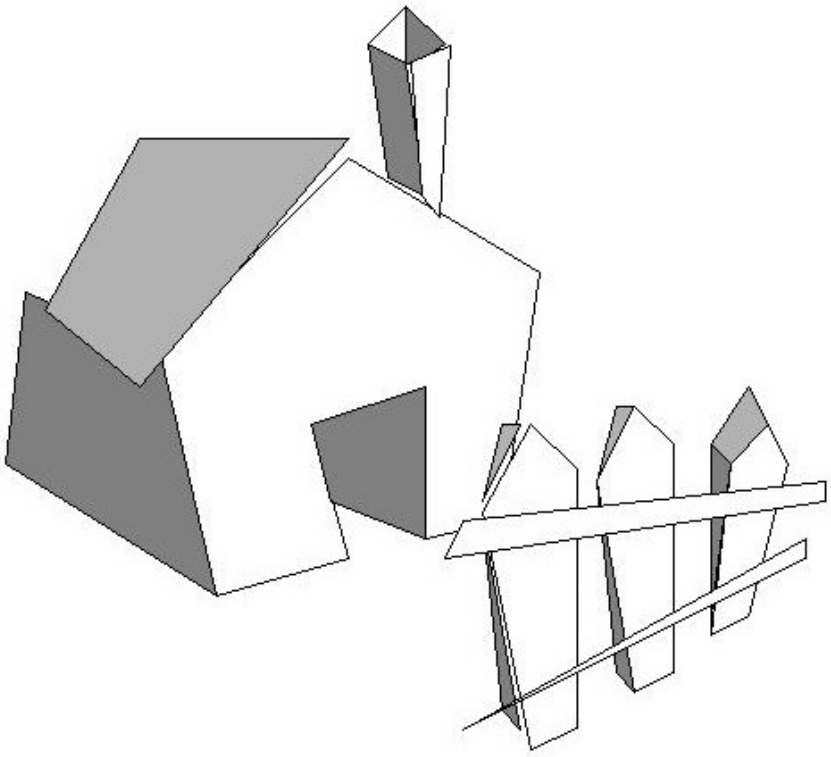
ABODE

Kernel Home Range Estimation for ArcGIS, using VBA and ArcObjects
By Pete Laver[1]

User Manual (Beta v.2 - 7 February, 2005)



[1]PETE N. LAVER, Department of Fisheries and Wildlife Sciences, Virginia Tech, 149 Cheatham Hall, Blacksburg, 24061-0321, 540.231.5320, plaver@vt.edu.

Table of Contents

1. Preface ***3***

2. Introduction ***3***

2.1. Home Range Analysis **3**

2.2. Software discrepancies **4**

3. Minimum Convex Polygons ***6***

3.1. The first home ranges **6**

3.2. Problems with polygons **7**

3.3. Benefits of polygons – Simple can be good **7**

3.4. A final note on comparability **9**

4. Kernel Density Estimation ***9***

4.1. The move from deterministic to probabilistic techniques **9**

4.2. What to make of all these user inputs **15**

4.3. Selecting a smoothing factor **16**
4.3.1. Non-statistical methods 16
4.3.2. Statistical methods 16

4.4. Discretization and the effect of rounding error **21**

4.5. Standardization **26**
4.5.1. Unit Variance Standardization 26
4.5.2. X Variance Standardization 28
4.5.3. Covariance Bias 29

5. Home Range Asymptotes ***30***

5.1. Why we should look at them **30**

5.2. How we should analyze them **31**

6. Core Home Ranges ***33***

6.1. Does a core really exist? **33**

6.2. How do we test for this? **33**

7. Data driven and Biologically meaningful methods ***36***

8. Using ABODE ***39***

8.1. How to start using ABODE **40**
8.1.1. Loading the form into the VBEditor 40
8.1.2. The VBA realm 42
8.1.3. The easy start-up for ABODE 45

8.2. Using ABODE for home range analysis **46**
8.2.1. The Visual Basic form and error trapping 46
8.2.2. Minimum Convex Polygons (MCPs) 48
8.2.3. Kernel Density Estimation 53

9. Conclusion ***59***

10. Acknowledgements ***59***

11. References ***60***

1. Preface

Home range analysis is an important part of our study of animals and it is fraught with problems from data collection to final management implementation. ABODE is user-friendly freeware that can be used in ArcGIS to do both MCP and kernel analyses. Both batch and single processing are available, as well as automatic functions for asymptote analyses and core home range analyses. The code for this package is open, and can be manipulated to suit the needs of the user with minimal VBA and/or ArcObjects experience. The functionality in and the detailed documentation provided with ABODE are aimed to address some of the more contentious issues in home range analysis.

Beyond providing documentation for ABODE, this 'user manual' is also aimed at giving the user background in home range theory and home range analyses. Should the user feel comfortable with the theory, they should skip to the documentation for ABODE, towards the end of the manual.

2. Introduction

2.1. Home Range Analysis

One of the basic requirements in the study of animals is an understanding of the relationship between the animal and its environment. At a grossly simplified level, this requirement is sufficed with home range analyses. Home range is a concept that attempts to describe the spatial context of an animal's behavior. Home Range was formally defined by Burt (1943): "… that area traversed by the individual in its normal activities of food gathering, mating, and caring for young. Occasional sallies outside the area, perhaps exploratory in nature, should not be considered part of the home range." From the outset, home range has been measured in terms of the hard boundary defining the edge of polygon containing the area used by an animal.

Home ranges have been analyzed since the earliest of hunting cultures first started to track their quarry. Today, techniques and uses of home range analysis have become more sophisticated. Home range theory is now used in conservation and management strategies. It is used to delineate protected area boundaries. As legal mandates for the protection of imperiled species have become more common (e.g. Endangered Species Act of 1973, USA), understanding the spatial requirements of an individual and then collectively for a minimum viable population has

become more important. The spatial context of animal behavior is important not only in the management of threatened species that we aim to preserve but also in the management of threats (e.g. invasive species). With the improvement in our conceptual grasp of home range, and our improvements in home range analysis, the concept of home range has been used for habitat evaluation and recently, home range theory has even been used to predict fitness (Mitchell and Powell, 2003).

Home range theory can be used for many conservation and management ends. Three things should dictate the methods used for doing this: the objective of the study; the potential of the data; and the life history of the animal. Unfortunately, there is no perfect home range estimator, and certainly no estimator should be used indiscriminately of the concerns listed above. Home range estimators should not be used as a black box, but should be used as they are intended – as tools to aid our improved understanding of animal behavior. The workmanship of the user rather than the tool will determine the quality of the final product.

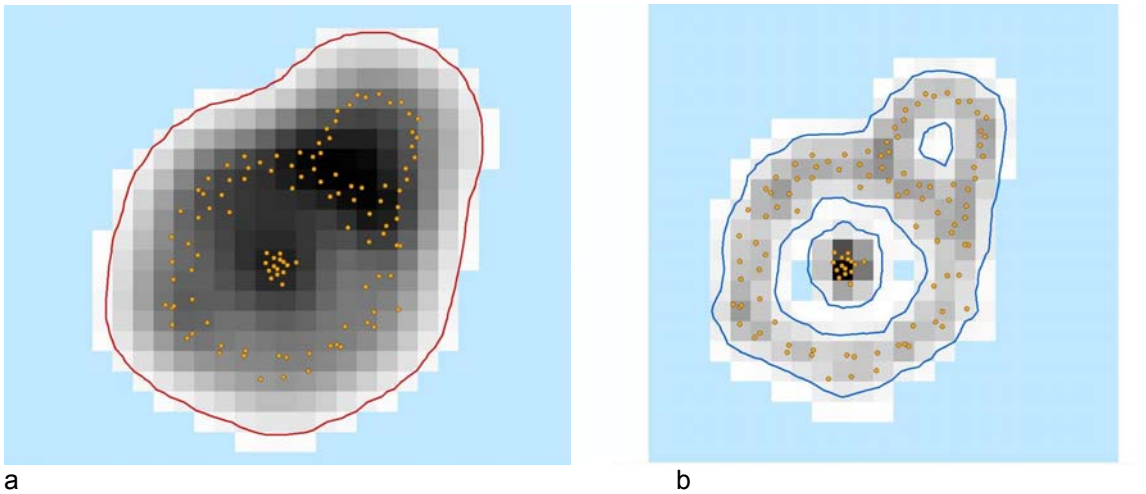
This document is intended to give some general background concerning home range theory as well as to describe one of the tools available for home range analysis, ABODE.

2.2. Software discrepancies

What software should I use to analyze home range? Larkin and Halkin (1994) reviewed several software packages used in estimating animal home ranges. At that point in time, few options were available for kernel estimation, and no comparison was possible. Lawson and Rodgers (1997) made a similar comparison using a single real dataset. They found significant differences between the results from the programs reviewed (CALHOME, RANGES IV, RANGES V, and TRACKER). The differences in kernel estimation were attributed to algorithms used in the programs. For this reason, it is recommended that algorithms be clearly stated in program documentation. They also noted that there were differences in the options that users were given in terms of the type of smoothing (fixed and adaptive) and the actual kernel used in the analysis. In the time since those publications, several other packages have become available. Most notably, two extensions for ArcView 3x have come into popular use. These are the Animal Movements Extension to ArcView v2.0 (AMAE)(Hooge and Eichenlaub, 2000), and the Home Range Extension (HRE)(Rodgers and Carr, 1998). Unfortunately no detailed review has been done for these kernel estimators. Several Minimum Convex Polygon (MCP) estimators are available, but it is hoped that different packages would provide identical results for a deterministic measure such as MCP.

They are both extensions that can be added easily to ArcView, a software package that has previously been the most commonly used package for ecological Geographic Information Systems. Many studies have reported the use of one or the other. Since ArcGIS 8x became available, many research and academic institutions have started to use this software as a replacement for ArcView. It was my goal to provide home range freeware that could be used in ArcGIS, which fully documented the algorithms and techniques used in analyses, and which would use the elements from the available ArcView extensions that were most helpful. I aimed to improve on certain analyses that I felt were not implemented optimally. Though many software discrepancies do exist, and no estimator (certainly not ABODE) will be perfect, it is suggested that the user read some of the following points highlighting differences that could have a considerable effect on a home range analysis.

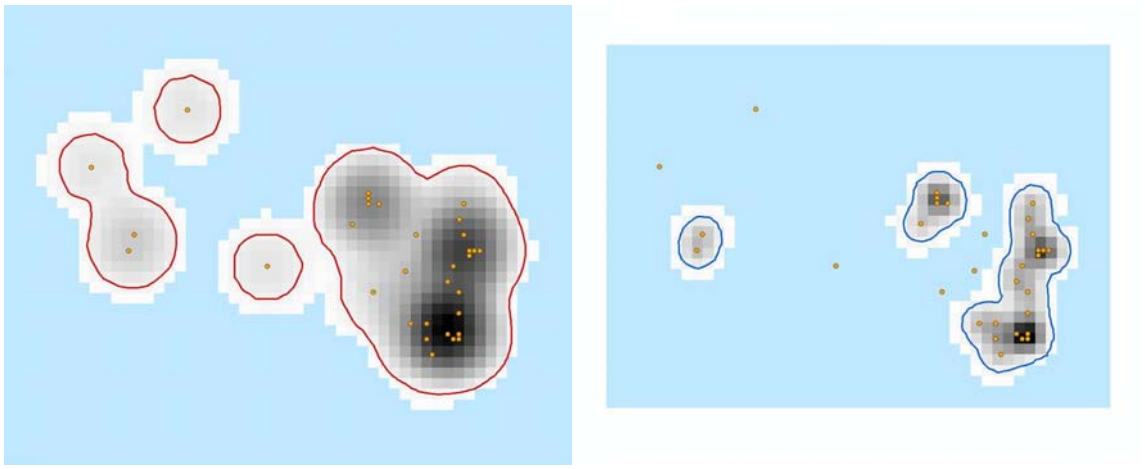
One discrepancy between currently favored home range estimators and ABODE is the ability to deal with 'islands' and 'donuts' (for a terrestrial species these may be features such as lakes or uninhabitable areas within the range). In Figure 2.2.1.a., a commonly used estimator is not able to eliminate areas of low probability within the distribution, as does ABODE (Figure 2.2.1.b.). In both figures, the user defined fixed smoothing factor was 32m, and both had a grid cell size of 15m. In ABODE, a biweight kernel was used, with a normal kernel being used in the other package. Standardization was not implemented in either package. Both figures are displayed at the same scale, and clearly there are not only differences in the extent of smoothing given the same inputs, but also in the ability to distinguish probability differences at a fine resolution.



a b

Figure 2.2.1. 95% Volume fixed kernel home ranges using no standardization, a normal (AMAE) or biweight (ABODE) kernel, h=32m, and 15m grid in software AMAE (red line), and ABODE (blue line). Orange points are a hypothetical set of location estimates. Blue cells in the grid indicate 100% probability (no density value), while dark cells indicate higher density than lighter cells.

The difference in software packages due to the contribution of outliers in a dataset is dealt with further in Section 4.1. Very briefly, a commonly used estimator assigns a density value to pixels that surround even outlying data. The effect is best observed when comparing Figures 2.2.2.a. and b., in which the former is the commonly used estimator, and the latter is the product of ABODE (for the same dataset). Both analyses had the same user inputs, which are discussed in Section 4.1. The inclusion of outliers (Burt's [1943] 'occasional sallies') results in glaring differences in home range size using kernel estimation.



a b

Figure 2.2.2. Single female cheetah dataset, analyzed using identical user inputs, in AMAE (red line) and ABODE (blue line). AMAE appears to assign values to outliers in the dataset and appears to assign value to cells further than the chosen smoothing parameter (distance from orange points to white grid cells).

This is one example of software discrepancies that result in different home range sizes and area use patterns. The are numerous others that the user should be aware of.

3. Minimum Convex Polygons

3.1. The first home ranges

Mohr (1947) introduced the concept of "minimum home ranges", and with it the idea of using a minimum convex polygon (MCP) to delineate a home range boundary. Since then, MCPs have been the most widely used home range estimation tool. Beyond the minimum convex polygon, the other major deterministic technique used in home range analysis is the grid cell count (Siniff and Tester, 1965).

3.2. Problems with polygons

The definitional and analytical flaws associated with the MCP estimator are reviewed elsewhere (Powell *et al*., 1997; Seaman, 1993; White and Garrott, 1990; van Winkle, 1975; Worton, 1987). Use of the minimum convex polygon encourages the notion of home range as a 2-dimensional entity with even space use (Powell, 2000). This is contrary to the cognitive map of variable resource and landscape value (Peters, 1978) that is manifested as a 3-dimensional entity of variable and potentially clumped use with a "diffuse and general" boundary (Stickel, 1954; Gautestad and Mysterud, 1993; Gautestad and Mysterud, 1995; Powell, 2000).

The MCP method has been shown to be highly sensitive to sample size (number of locations) (Seaman *et al*., 1999; Bekoff and Mech, 1984; Laundré and Keller, 1984; Harris *et al*., 1990; White and Garrott, 1990; Kernohan *et al*., 2001). This issue is dealt with in the next section (3.3.) as well as in the discussion pertaining to home range asymptote analyses (Section 5). The sample size issue is also related to the inability of MCP's to objectively treat outliers (Seaman *et al*., 1999). Further problems include the sensitivity of the estimator to spatial resolution (Hansteen *et al*., 1997), and sampling duration (Swihart and Slade, 1985a; Powell, 2000).

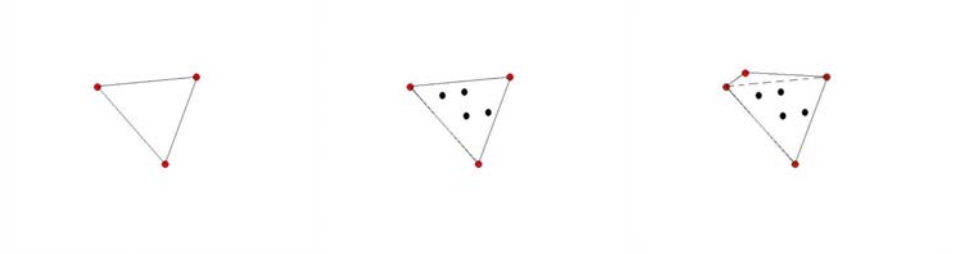
Most studies cite the use of MCPs for reasons of comparison. Unfortunately this notion only engenders inappropriate comparison because of the sample size issues. Only the most meticulously matched studies should use MCPs as a form of comparison, if the sample sizes are equivalent, along with equal sampling durations and similar treatment of outliers. Having stated this, MCPs do have a place in home range analysis.

3.3. Benefits of polygons – Simple can be good

Choice of home range estimator should depend on three factors (see above), namely the objective of the study, the nature of the data, and the movement behavior of the animal in question. Minimum convex polygons do have a place in home range estimation where these three factors are satisfactorily incorporated into the choice of MCP. Sometimes, the objective of the study is only to find the entire area used by an animal, even if this does not meet the commonly held definition a home range given by Burt (1943). In such cases, analyses should not be reported as home range estimates, but rather as total range estimates. These are often as important to managers and conservationists as are home range estimates. They could indicate sensitivity to external threats that may only be contacted once in a lifetime, but that could be deleterious to the animal. A common threat of this sort would be becoming a pathogen vector or invasive species vector. An example of this could be the contraction of an infectious disease at

the edge of a range or perhaps persecution upon leaving a protected area (e.g. *Lycaon pictus*, Woodroffe and Ginsberg, 1998). MCPs might highlight this sort of high-risk (but infrequent) movement behavior. Very often, the data used in home range analyses are so poor that MCP might be the best available science. Seaman *et al*. (1999) showed that relatively large sample sizes (number of locations per animal) are required for kernel analyses. In cases where sample sizes are insufficient, an MCP estimate of area used (though not necessarily a home range) will be better than nothing. Occasionally, the movement behavior of an animal is such that the uniform area-use suggestive of MCPs may in fact be valid. This would be the case if the animal's location estimates showed an even distribution. In such cases, kernel analyses may not provide much more detail than an MCP since the density of locations will be equal everywhere in the interior of the area of use.

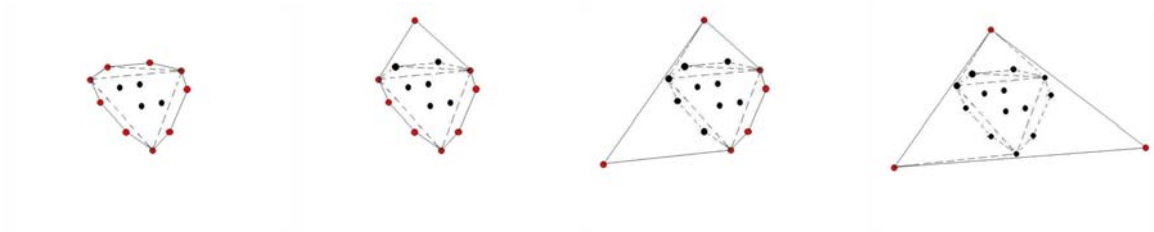
The sensitivity of MCP estimates to sample size (and sampling duration when points are added sequentially), and outliers, may be understood in the simplified sequence of Figure 3.3.1. and Figure 3.3.2. These sequences show how the outliers that might usually make MCP inappropriate for analyses, may indicate important biological events such as dispersal - events that are important to our understanding of an animal's natural history and to its conservation (Woodroffe and Ginsberg, 1998). The simplest polygon is a triangle of three points (Figure 3.3.1.a.). As the animal moves around, points may be added within the boundary formed by that triangle (Figure 3.3.1.b.). Eventually the animal will extend the estimate of area used by moving outside of the perceived boundary (Figure 3.3.1.c.).



a b c

Figure 3.3.1. Sequence of hypothetical location estimates added sequentially, with resultant increases in MCP home range estimates (polygons).

As this continues (and as sample size increases) the estimate of area use increases (Figure 3.3.2.a.). Eventually certain exploratory movements (sallies) or directed dispersal events will greatly increase the estimate of area used (Figures 3.3.2.b-d). These sallies can be evaluated most easily through the use of a home range asymptote analysis (Section 5).



a b c d

Figure 3.3.2. As exploratory movements are added to the dataset, the polygon defining the home range boundary increases. The addition of a few "sallies" can greatly increase the area of the polygon. The decrease in the number of points defining the polygon boundary (red points) may indicate the addition of exploratory movements.

3.4. A final note on comparability

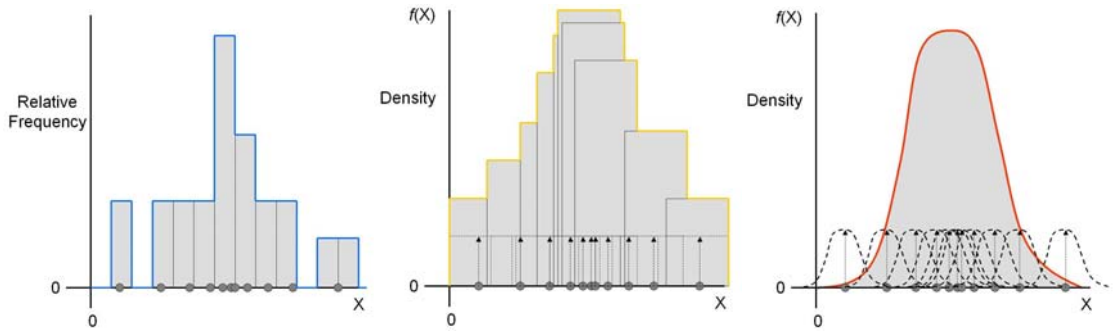
While there is still a place for MCPs in home range analyses, usually the polygon will not define the home range, but rather the entire area used by the animal. This is a useful piece of information in its own right, but should not be compared across studies, as is so often the case in literature. If the MCP is reported for comparison's sake then the sample size, sampling duration and treatment of outliers should be explicitly stated (and hence matched) for each individual analyzed.

4. Kernel Density Estimation

4.1. The move from deterministic to probabilistic techniques

We rarely have complete information about an animal's movement behavior. The result of this is that deterministic techniques will be heavily biased by our sampling methods. This may also be true for probabilistic techniques. The latter use interpolation and extrapolation based on the distribution of the data. They provide a more robust analysis that acknowledges the importance of the distribution of the data as a whole, rather than evaluating each point in isolation. The first formal probabilistic techniques were circle (Hayne 1949) and ellipse (Jenrich and Turner, 1969; Dunn and Gipson, 1977) approaches. Both of these techniques are parametric, assuming a bivariate normal distribution of the location data – an assumption that is generally violated (Powell, 2000). Nonparametric estimators in this group include Fourier series (Anderson, 1982), harmonic mean distribution (Dixon and Chapman, 1980), and kernel estimators (Worton 1987, Worton, 1989). Kernel estimation is currently the most widely used home range technique.

Silverman (1986) first described kernel density estimation for the layperson. As a result, much of the subsequent kernel home range literature rests heavily on this work. Silverman describes kernel estimation as follows (Figure 4.1.1. adapted from Silverman, 1986). Density estimation in its simplest form is a histogram representation of the variable in question (i.e. the x or y coordinate of a set of locations) (Figure 4.1.1.a.). The naïve estimator is a histogram constructed such that each point falls in the centre of a sampling interval (bin) (Figure 4.1.1.b.). In this method, the sampling intervals overlap, and the points that are included in any interval are weighted according to a uniform distribution. The kernel estimator is an improvement of this naïve estimator in that it replaces the uniform weighting function with a kernel function. This kernel function is a probability density function with a specified distribution (Figure 4.1.1.c.). Equation 1 defines the kernel density estimator:



a b c

Figure 4.1.1. Progression of density estimation from simple histogram techniques to smooth kernel estimation (Adapted from Silverman, 1986). In these examples, the abcissa (x axis) could be x coordinate for a set of locations, that in the animal spends more time or is seen more in the center of the distribution in the x direction.

$$\hat{f}(x) = \left[\frac{1}{(nh^2)}\right]\sum_{i=1}^{n} K\left\{\frac{(x - X_i)}{h}\right\} \qquad (1)$$

K is the kernel that determines the shape of the distribution that is placed over each of the points. *h* is the smoothing factor (also known as the smoothing parameter, bandwidth and window width), which controls the search radius or width or the kernel. *n* is the number of location estimates (points) used in the analysis. *x* and *X* refer to the vectors of the coordinates of the evaluation point and all other points, respectively. Various kernels have been described (Silverman, 1986), but kernel analysis has been shown to be relatively insensitive to the choice of kernel (Silverman, 1986). Seaman and Powell (1996) use, as an example, the Biweight Kernel, K_2 from Silverman (1986). Equation 2 defines the biweight kernel:

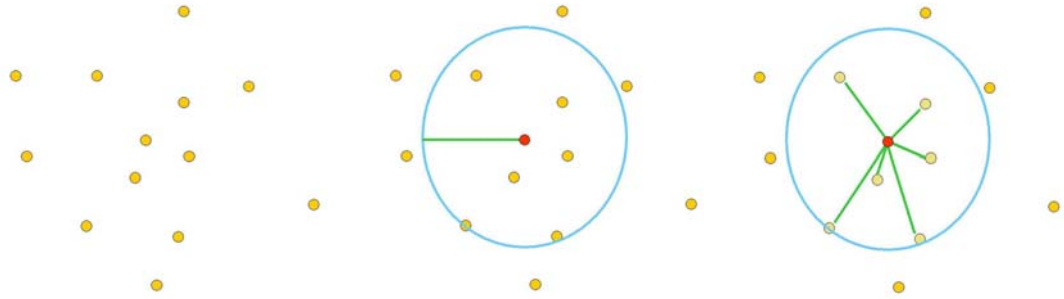
$$K_2(x) = \begin{cases} 3\pi^{-1}(1 - x'x)^2 & if \quad x'x < 1 \\ 0 & otherwise \end{cases} \qquad (2)$$

x'*x* is the distance from the evaluation point to any other point in the set, divided by the smoothing factor, h. Thus, if *x*'*x* < 1, then the point in question is within the search radius (*h*) of the evaluation point and is used in estimating the density at the evaluation point. If *x*'*x* > 1 then the point is too far away from the evaluation point to be considered. Once a point is included in the density estimation, its contribution is weighted using an inverse distance weighting function (the kernel). This makes intuitive sense, since a point that is near the periphery of the search area will have a large distance (*x*'*x* tends towards 1) and should consequently contribute less to the density estimate than a point close to the evaluation point. This kernel is calculated more quickly than the normal kernel, and has higher differentiability properties than the Epanechnikov kernel, two of the other commonly used kernel options that are available (Silverman, 1986). The final probability density function (pdf) will reflect the shape of the underlying kernel, and thus a pdf based on a biweight kernel will tend to be smoother than that based on an Epanechnikov kernel.

In the following paragraphs, I hope to provide a functional description of how kernel estimation may be done. ABODE uses the following methodology in its kernel estimation. This process is a simplification of a true kernel density estimate, but provides a reasonable proxy. Ideally, a kernel function is placed over each point in the dataset. Where there are other points in the vicinity (within the search radius or *h*) then the kernel has a value that reflects the density at that particular point. Where the point being evaluated is isolated, the kernel value is zero (if a biweight kernel is used). Ideally, a grid of infinitesimal resolution is placed over the distribution of points. At each grid intersection (or from the center of each grid cell) the values of each of the kernels are summed. This provides a continuous surface depicting the probability density function. In reality, this analysis would take an almost infinite amount of time. We thus select a grid size to represent the most favorable tradeoff between resolution (and hence smoothness of the pdf) and time (course grids taking less time to analyze). Even with a reduced number of pixels to evaluate, the process is still time consuming. A short cut can be taken by selecting only the pixels that have a chance of having a value greater than zero – i.e. only those within the search radius of those points that have not (effectively) been eliminated by the biweight kernel.

In kernel estimation, each point in a given distribution (Figure 4.1.2.a.) is evaluated. Each evaluation point (red in Figure 4.1.2.b.) is in turn evaluated based on the points that surround it. A point that is surrounded by many other points will have a high density value. To determine which surrounding points will contribute to the estimation of the density at the evaluation point, a smoothing factor (bandwidth), h, is used to describe the search radius about the evaluation point

(green radius of the blue search area in Figure 4.1.2.b.). The distance from each point (within the search radius) to the evaluation point is then calculated (Figure 4.1.2.c.). Based on these distances, a cumulative value is assigned to the evaluation point. Next, another evaluation point is selected (Figure 4.1.3.a.). This procedure continues until all the points in the distribution have been evaluated. They are all scored and assigned density values (denoted by classified symbology in Figure 4.1.3.b.). A grid of specified size is then overlaid on the distribution (Figure 4.1.3.c.). Starting again with each evaluation point (red), the pixels within the search radius are populated with (assigned) their respective density values (Figures 4.1.4.a. and 4.1.4.b.). Each subsequent point is evaluated in the distribution (Figure 4.1.4.c.). Thus two processes are occuring: first a point to point evaluation and then a pixel to point evaluation.

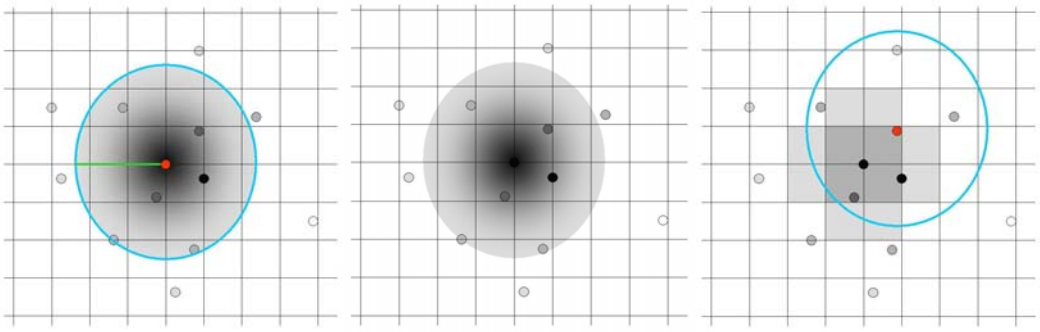


a b c

Figure 4.1.2. Process of finding points in a distribution (yellow points) that will contribute to the density estimate at an evaluation point (red point). The area searched (blue circle) is determined by the smoothing parameter, h (green line).

a b c

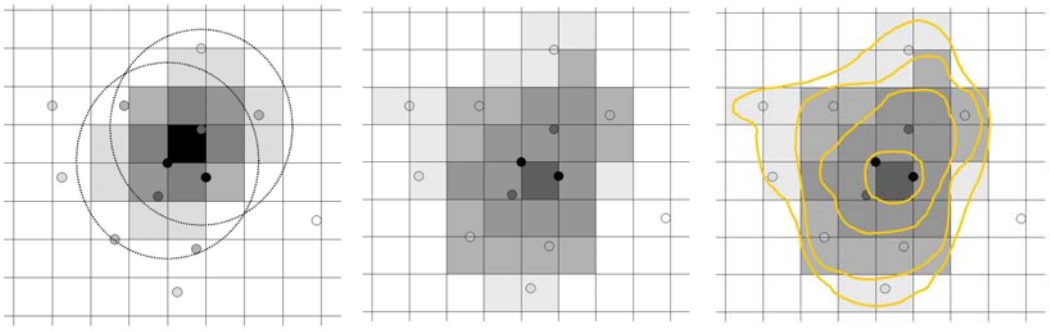
Figure 4.1.3. The procedure for generating density estimates continues from point to point in the distribution (yellow points) until all locations have a density value (high density in black, to low density in grey; white indicates zero density). Overlaying a grid is the next step towards building a density surface.



a b c

Figure 4.1.4. Assigning pixel density. Pixels or grid cells are given density values based on their proximity to an evaluation point. This happens for each location in the set.

Where the search areas for two evaluation points overlap, the density value assigned to a pixel is calculated cumulatively (Figure 4.1.5.a.). This is the same as summing the value for the kernel at every point on the surface. Finally a surface is created that contains pixel values of the kernel density estimate of the distribution (Figure 4.1.5.b.). The surface is then contoured at specified volumes to give percentage home ranges (i.e. a 95% home range is contoured at 95% of the volume of the density surface – not at 95% of the area of the home range) (Figures 4.1.5.c.).

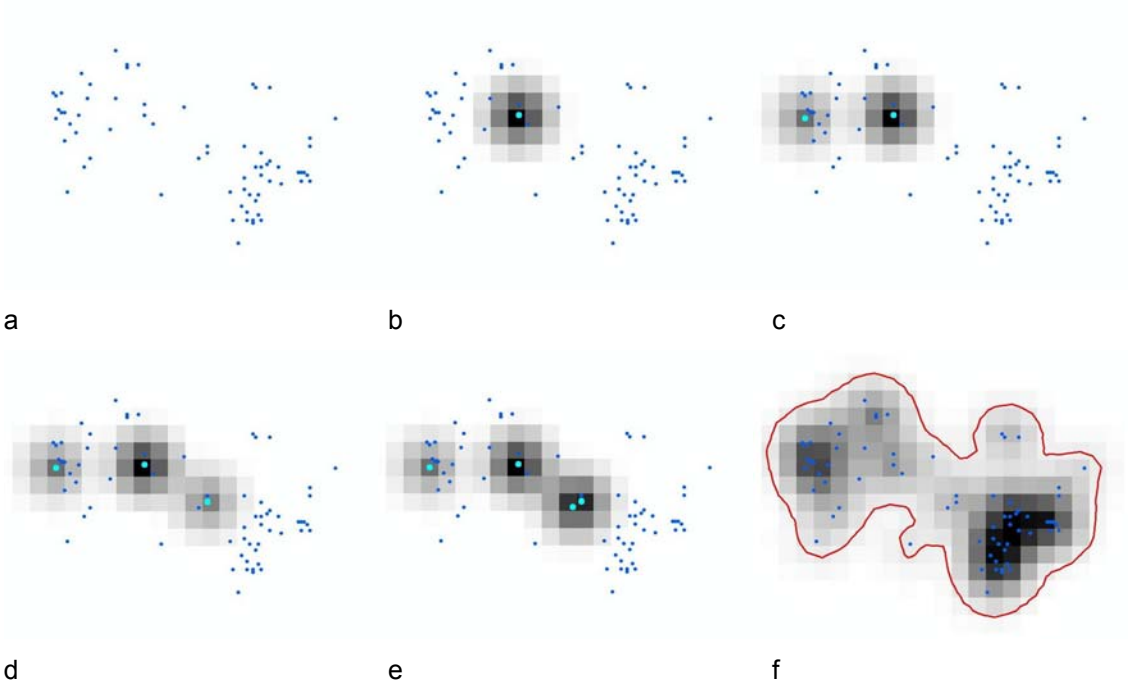


a b c

Figure 4.1.5. Pixel density accumulation (values assigned to pixels from different evaluation points are summed) and volume contouring (defining the 2-dimensional area that contains a specified percentage volume of the density surface).

To see this process in effect for a real dataset, Figure 4.1.6 shows the progression of pixel evaluation for the locations of a female cheetah in the Serengeti National Park, Tanzania. The second point in the dataset was at the same location as the first point. The evaluation for this second point was omitted from the sequence, but the result can be seen in the evaluation for point 3 (Figure 4.1.6.c.), where the darker cells for points 1 and 2 indicate a higher density (where the pixel values were summed). In this sequence, the final contour shows the 95% (volume) home range for a user-defined 300m grid and a user-defined smoothing factor of h = 1000m,

using a fixed biweight kernel and no standardization, in ABODE (Figure 4.1.6.f.) (standardization is explained in Section 4.5).



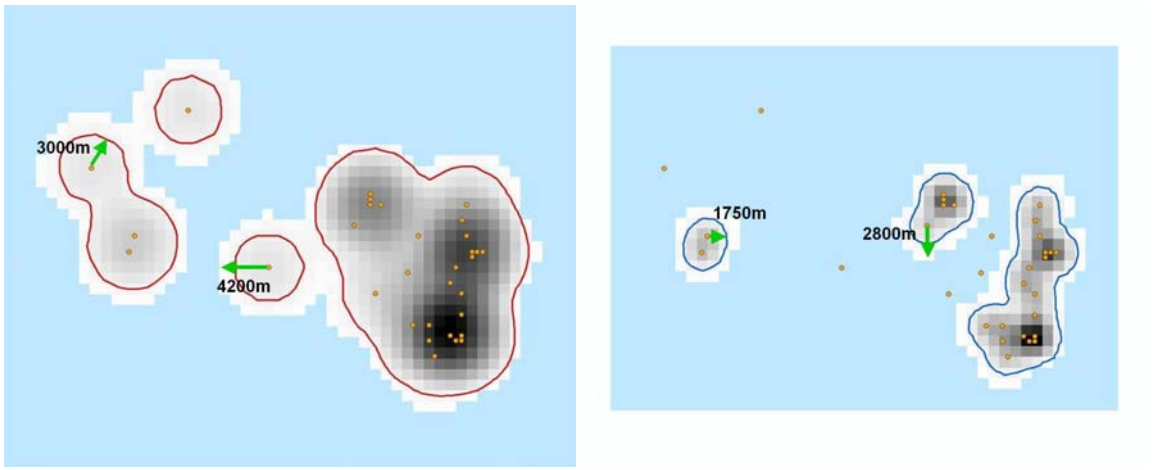
a b c

d e f

Figure 4.1.6. Progression of kernel density estimation for a real dataset.

This method of point-to-point evaluation, followed by pixel to point evaluation is a short cut, but serves the same purpose as the more lengthy analysis at each and every pixel. The added benefit of this procedure is that outliers in the dataset are effectively eliminated from the analysis before the time consuming pixel to point analysis begins. Using a biweight kernel satisfies one of the tenets espoused by Burt (1943) in his home range definition. In this case the occasional sallies are essentially removed from the dataset, though it must be understood that this discrimination is based solely on the spatial qualities of the distribution, and not on verified dispersal or exploratory movements. It is important that the user understands how these outliers are treated in different software packages. In one of the most commonly used packages available at the moment (Figure 4.1.7.a.), pixels surrounding outliers are giving a density value, albeit a very low value. When contouring at 95% of the volume of the density surface, this may not be a problem, since the outliers will probably not be picked up. No testing has been done, to see what the likelihood is of seeing significant differences at different percentage home ranges (volume contouring). In the example shown, the commonly used estimator (AMAE) evaluates pixels up to 4200m away, when the smoothing factor selected was 3000m. The contours begin at 3000m from the points. In Figure 4.1.7.b., ABODE evaluates only pixels for which their center is within 3000m from the points, and only for those points that have other points within the search radius. In this case, contours depend on the volume of the density surface, and may begin for

example at only 1750m from the points. Both Figure 4.1.7.a. and Figure 4.1.7.b. are at the same scale – the difference in size of the bounding box shows the extent of the grid used in each (with blue pixels showing a density value of 0). Both analyses used fixed kernels, with a user defined smoothing factor of 3000m and 1000m grid cells. Data were not standardized in ABODE, and a biweight kernel was used. In AMAE no standardization was used and a normal kernel was used instead of the biweight kernel.



a b

Figure 4.1.7. Comparison of two kernel packages, AMAE and ABODE using the same user inputs of a 3000m smoothing factor (h) with 1000m grid cells. The only apparent difference in analysis is the use of a normal kernel in AMAE (a) and the use of a biweight kernel in ABODE (b), though this should not result in the differences obtained (see Section 4.2).

4.2. What to make of all these user inputs

Running a kernel analysis can be a daunting task considering all the user inputs that are required. To date there is no 'industry standard' for the protocol to be followed when estimating a kernel density. This can be both good and bad. It is good since no single protocol will be suitable or desirable in all situations. It is bad since the estimates produced for a kernel analysis will only strictly be comparable if the same protocol is followed (i.e. the inputs are the same). It is evident from various sources (Epanechnikov, 1969; Silverman, 1986; Worton, 1987) that the kernel used (i.e. biweight or normal) will have little effect on the outcome of the analysis, but the nature of the kernel, fixed versus adaptive, and the smoothing factor especially will greatly influence results (Silverman, 1986). A fixed kernel uses a single smoothing factor (h) for evaluation of the entire dataset, whereas an adaptive kernel evaluates a new smoothing factor for areas of the dataset with different densities. It is unlikely that a consensus will be reached in the near future about which inputs should be used. For this reason, **the user should always state all user inputs when reporting kernel results**. If these are reported, then studies with the objective of

comparison may be tailored to use the same estimator inputs ensuring that the studies are comparable.

4.3. Selecting a smoothing factor

As stated above, the most important input when doing kernel density estimates is the smoothing factor (Silverman, 1986). This decision requires some consideration of the nature of the data and of the behavior of the animal. Seaman and Powell (1996) suggested the use of fixed kernels with Least-Squares Cross-Validation as the technique for choosing the smoothing parameter.

4.3.1. Non-statistical methods

Non-statistical methods are usually not as robust as statistical methods, in the sense that they are more subjective and are not necessarily always reproducible. The starting point for any kernel analysis should be to determine if there are any factors that may guide the selection of a smoothing parameter. These factors may relate to either the natural history of the animal, or to the nature of location data. An example of a method that incorporates these two concepts is detailed later in section 7. If there are no insights available to guide the choice of smoothing parameter, then the user is left with two options. The user can subjectively choose the appropriate smoothing factor based on a candidate set of curves (i.e. run the the analysis using various smoothing factors, and choose the option the resulted in the 'best' fit). The smoothing parameter used to obtain the curve that best suits the objectives of the analysis should then be used. While this is not a statistically defensible choice of h, it may allow for a greater range of objectives in the analysis. This type of method may allow the user to elucidate fine scale detail in the distribution, or to get a very general description of the overall pattern of density, and any level of detail necessary within this range. The second option would be to use a more statistically defensible and repeatable method, thus automatic methods for choosing the smoothing parameters may be favored. These are described in the next section.

4.3.2. Statistical methods

Two statistical methods that are commonly used for automatically choosing a smoothing parameter are the reference smoothing method and the least-squares cross validation method.

Reference Smoothing

Silverman (1986) suggested that an optimum bandwidth could be chosen for a standard distribution – the bivariate normal distribution. Though this method is known to oversmooth data that are multimodal and non-normal, it does provide an easy and automatic starting point with which to analyze datasets. Following Silverman (1986), the optimum smoothing parameter for a bivariate normal density is:

$$h_{opt} = \frac{\sigma}{\sqrt[6]{n}} \quad (3)$$

Worton (1995) referred to this optimum bandwidth as h_{ref}. n is the sample size. The standard deviation term is calculated as (Worton, 1995):

$$\sigma = \sqrt{\frac{\sigma_x^2 + \sigma_y^2}{2}} \quad (4)$$

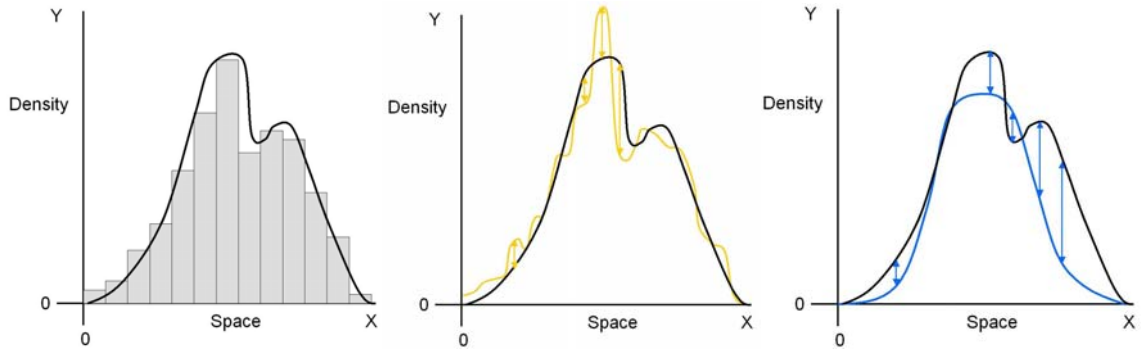
σ_x^2 and σ_y^2 are the variance estimates for the data in the x and y directions. Since most animal movement data aren't normally distributed, h_{ref} will seldom provide a good estimate of home range size, and will generally overestimate. In ABODE, a biweight kernel is used, and the reference smoothing factor should be adjusted using a constant *A(K)* (see Equation 13 and supporting text).

Least-squares Cross Validation

Least-squares cross-validation (LSCV) was proposed by Rudemo (1982) and Bowman (1984) as another automatic method for selecting the optimum bandwidth or smoothing parameter based on the unknown density of the given distribution of points. This method involves the minimization of a loss function. The integrated squared error loss function was suggested by Bowman (1984). Bowman (1984) showed that the integrated squared error loss function performed better than the Kullback-Leibler loss function proposed by Rudemo (1982) for distributions with long tails (i.e. those that deviated from normality). Bowman (1985) showed that this squared-error cross-validation technique performed consistently better than several other loss functions. Usually, this process would require integration, but by using a normal kernel, the evaluation may be done analytically (Bowman, 1984). The normal kernel is computationally simpler to use in least-squares cross-validation, but is not necessarily used in the actual density estimation. The constant *A(K)*, can be used to modify the optimum smoothing factor chosen by LSCV, when estimating density using other kernels (Silverman, 1986). Seaman and Powell (1996) followed similar methodology and used the constant *A(K)* = 2.78 for the conversion from a normal to a

biweight kernel (this is an unexplained discrepancy between Seaman and Powell, 1996 and Silverman, 1986).

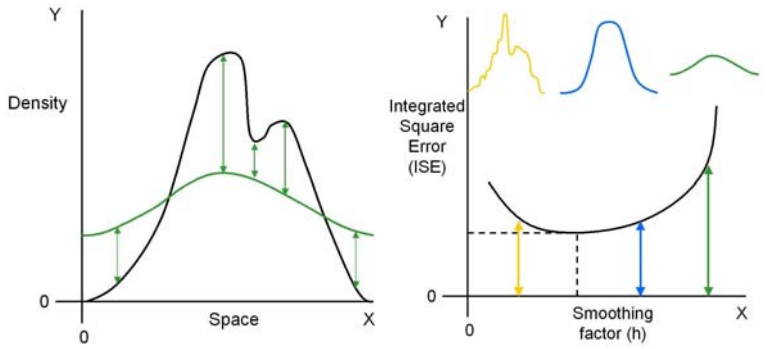
Given a true density function for some distribution of data (Figure 4.3.2.1.a.), various values of *h* (the smoothing parameter) are used to obtain density estimates. For example, Figure 4.3.2.1.b. shows the density estimate given a smoothing parameter smaller than the optimum (or true) value. The difference between the estimated and true density is evaluated as the sum of the area of deviation of the estimate from the true density. Figure 4.3.2.1.c. shows that difference given a smoothing parameter slightly larger than the optimum value.



a b c

Figure 4.3.2.1. Least-Squares Cross-Validation with the true density (black line) (a) and an over- (yellow line) (b) and under-estimate (blue line) (c) of the smoothing parameter.

Finally, with highly oversmoothed data, as with a smoothing parameter that is considerably larger than optimum (Figure 4.3.2.2.a.), the difference between the density estimate and the true density is large. The loss function can be considered to be the difference in area between the estimate and the truth. In this case it is the area difference, but in home range estimation, with bivariate data, the loss function would be the difference in volume between the two surfaces. In reality this loss function is the integrated square error (integrated since we are dealing with a density, and square error, since we want to incorporate error in both over- and under-estimation). It is intuitive that we want the smallest deviation from the truth, and thus, when we plot the integrated square error for various smoothing parameters (Figure 4.3.2.2.b.), we search for the absolute minimum on the curve. At this point we find the associated smoothing parameter, which becomes our estimate of the optimum smoothing parameter.



a b

Figure 4.3.2.2. Generation of a loss function from LSCV.

The least-squares cross-validation score is $M_1(h)$, and is given by equation 5 (the definitions of X_i and X_j follow from Equation 1) (Silverman, 1986):

$$M_1(h) = \frac{\sum_i \sum_j K^* \left\{ \frac{(X_i - X_j)}{h} \right\}}{n^2 h^2} + \frac{2K(0)}{nh^2} \tag{5}$$

K^* is defined by

$$K^*(t) = K^{(2)}(t) - 2K(t) \tag{6}$$

Using the symmetric kernel for bivariate normal density (Gaussian density), $K^{(2)}$ is the convolution (multiplication) of the kernel with itself. This amounts to an equivalent kernel with a variance of 2 (Silverman, 1986). This is demonstrated by multiplying two bivariate normal densities, *f* and *g*, in equations 7 and 8 respectively to give *f*g* in equation 9:

$$f = \frac{e^{\frac{-(t-\mu_1)^2}{2\sigma_1^2}}}{\left(\sigma_1\sqrt{2\pi}\right)} \tag{7}$$

$$g = \frac{e^{\frac{-(t-\mu_1)^2}{2\sigma_1^2}}}{\left(\sigma_1\sqrt{2\pi}\right)} \tag{8}$$

$$f * g = \frac{e^{\frac{-[t-(\mu_1+\mu_2)]^2}{\left[2\left(\sigma_1^2+\sigma_2^2\right)\right]}}}{\sqrt{2\pi\left(\sigma_1^2+\sigma_2^2\right)}} \tag{9}$$

The multivariate normal density function, as described by Silverman (1986) is:

$$K(x) = \frac{e^{\left(\frac{-x'x}{2}\right)}}{2\pi} \tag{10}$$

Where $x'x$ is the distance between the evaluation point and another point in the distribution, divided by the smoothing parameter ($x'x$ = distance^2/h^2). Given equation 9, the convolution of the multivariate normal density function, $K^{(2)}$, becomes:

$$K^{(2)}(x) = \frac{e^{\left(\frac{-x'x}{4}\right)}}{4\pi} \tag{11}$$

The least-squares cross-validation score now becomes (Worton, 1995; Rodgers and Carr, 1998):

$$M_1(h) = \frac{\sum_i \sum_j \left(\left(\frac{e^{\left(\frac{-x'x}{4}\right)}}{4\pi} \right) - 2 \left(\frac{e^{\left(\frac{-x'x}{2}\right)}}{2\pi} \right) \right)}{n^2 h^2} + \frac{1}{\pi n h^2} \tag{12}$$

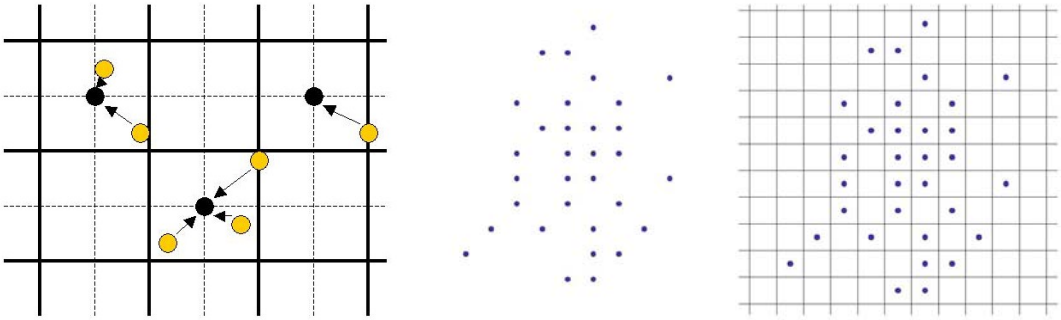
ABODE uses the minimization process, "Routine GOLDEN" (Sprott, 1991) to minimize $M_1(h)$, and hence find the optimum value for h for a bivariate normal density. The seed value used in the minimization process is the h_{ref} value (Equation 3), which is subsequently multiplied by 0.01 to get the lower bound of the search values. For unit variance standardization, h_{ref} is calculated from the standardized dataset (i.e. σ = 1). When it is to be used as the smoothing parameter, the h_{ref} value is adjusted for the biweight kernel. $A(K)$ is the constant used to make the adjustment. For conversion to a biweight kernel, A(K) = 2.04 (Silverman, 1986). As stated earlier, this discrepancy is unexplained.

$$h_{opt} = \frac{A(K)\sigma}{\sqrt[6]{n}} \tag{13}$$

Once an h has been estimated from LSCV, this value is adjusted for the biweight kernel by multiplying by the constant $A(K)$ = 2.04 (for use as a biweight kernel in the actual density estimation) (Seaman and Powell, 1996; Silverman, 1986:87). It should be noted that ABODE follows Silverman (1986) by using 2.04 as the constant, and not the 2.78 reported by Seaman and Powell (1996).

4.4. Discretization and the effect of rounding error

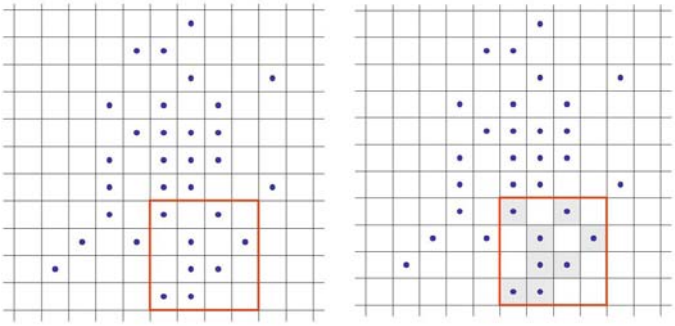
In many cases, location data for study animals are discretized or rounded. One example of how this may occur is when data are recorded using a map with a grid, such that location estimates are placed in grid cells (Figure 4.4.1.a.). In this case, many locations that fall within a cell may be rounded to the same point within that cell, resulting in considerable overlap of points. Datasets that display a regular pattern may fall into such a category (Figure 4.4.1.b. and Figure 4.4.1.c.).



a b c

Figure 4.4.1. Discretization in a dataset.

This form of discretization is an artifact of data recording. Sometimes the data collection scheme will also lead to this phenomenon, as with the use of a regular trapping grid. This is the case in small mammal trapping, regular line transects, and camera trapping grids. Here location estimates show the actual location, but may still result in considerable overlap since the sampling effort is concentrated in those specific places. With a discretized dataset (Figure 4.4.2.a.) a simple density estimate would show uniform density for the cells with points (Figure 4.4.2.b.).



a b

Figure 4.4.2. Density evaluation without taking into account overlapping points.

One could consider the same evaluation area in 3 dimensions, such that the x and y coordinates are plotted in the horizontal plane (Figure 4.4.3.a.), and the number of points is plotted on the vertical axis (Figure 4.4.3.b.). The number of points could (very simplistically) be a proxy for a density estimate (Figure 4.4.3.c.). When the overlapping points are displayed on the vertical axis (Figure 4.4.3.d.), the original density (Figure 4.4.3.e.) changes to reflect the clumping of points (Figure 4.4.3.f.). The final density estimate is not intuitive when compared to the 2-dimensional depiction of the distribution as is most common with our data (Figure 4.4.4.).

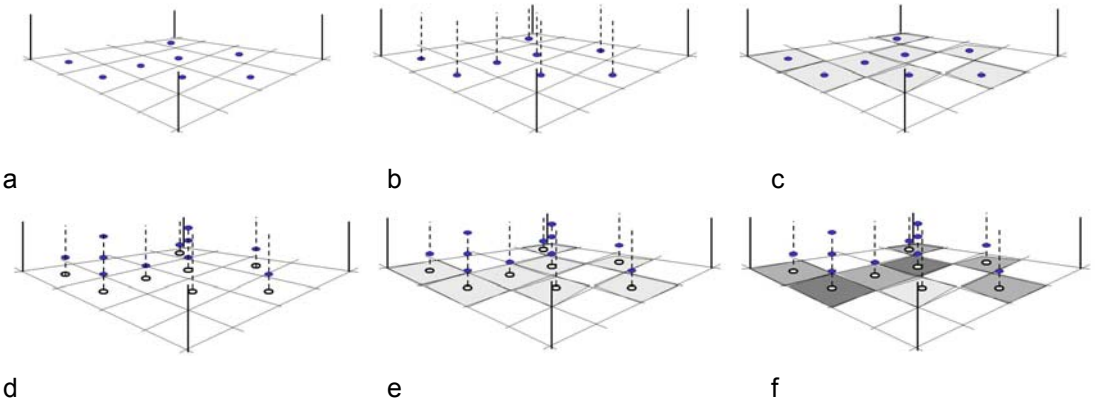


Figure 4.4.3. Demonstration of the effect overlapping points may have on density estimates, using simple grid cell counts.

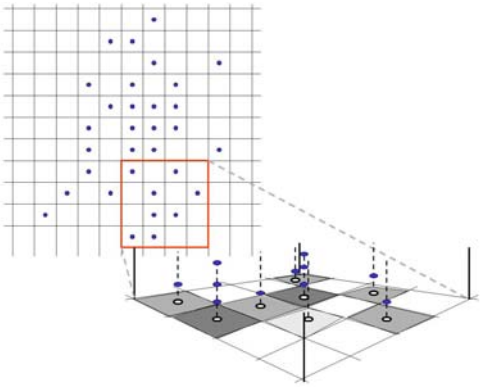
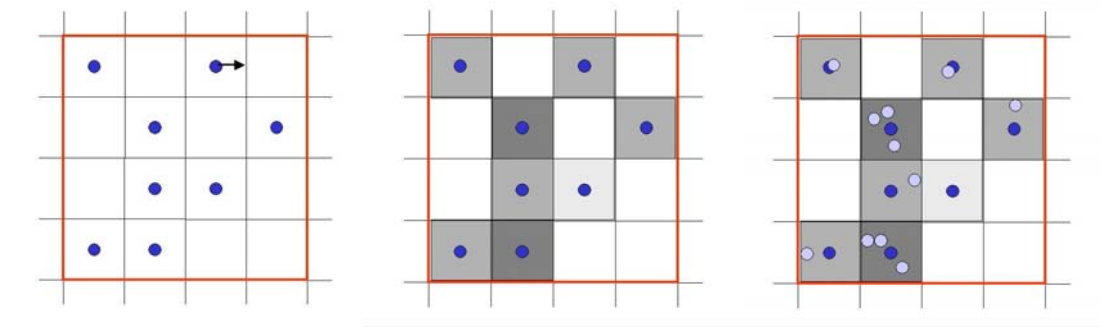


Figure 4.4.4. Comparison of 2D and 3D depictions of discretized data.

One method for depicting the overlapping distribution in 2 dimensions that would allow for improved interpretation would be to shift the overlapping points such that they would reflect the rounding error. This rounding error is roughly half of the diameter of the grid cell (Figure 4.4.5.a.), though greater error would occur if the initial location estimate was in the corner of a cell. By shifting the location estimates a randomly selected distance (between 0 and the rounding error) the density of the distribution (Figure 4.4.5.b.) may be better visualized (Figure 4.4.5.c.). In this case, depiction in 3 dimensions (Figure 4.4.6.) would not be necessary for density interpretation.



a b c

Figure 4.4.5. Visualization of density estimates produced by shifting location estimates when overlap occurs.

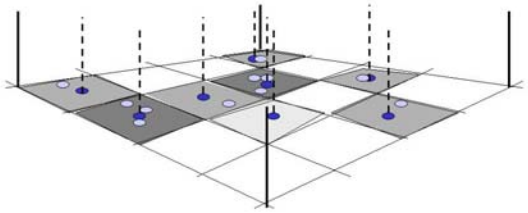


Figure 4.4.6. With shifting of overlapping points, 2D depiction is sufficient for interpreting density.

The real problem with discretization lies in the method used for smoothing parameter selection. Silverman (1986) concluded that in datasets with pairs of overlapping points $X_i = X_j$, least-squares cross-validation may show a minimization score tending towards infinity, and will choose a smoothing parameter that tends to zero. Where discretization is severe and overlap between points is extensive, LSCV may choose $h = 0$ for the smoothing factor. Since LSCV is the currently preferred automatic method for an objective choice of smoothing parameter, this may pose a problem in many analyses. Depending on the magnitude of the discretization and the extent of overlap, LSCV may underestimate the smoothing parameter to varying degrees, and may not always degenerate to $h = 0$. A clue that discretization may be playing a role in your data analysis would be the depiction of contours around single cells or small clusters (Figure 4.4.7.). Discretization effects may be borne out in contours that contract to form small islands throughout the distribution (Figure 4.4.7.).

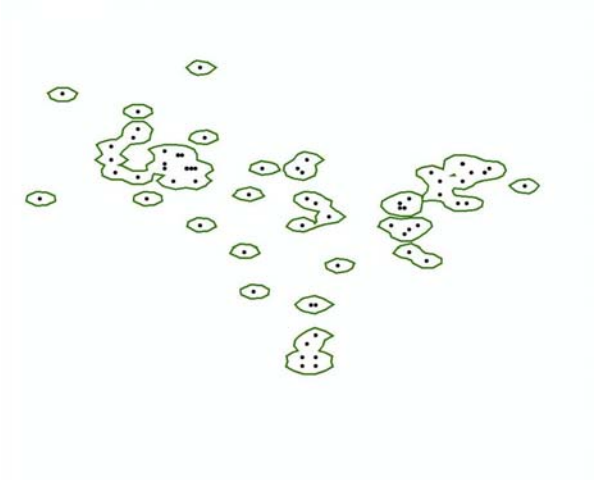


Figure 4.4.7. Effect of discretization (and overlap) in a female cheetah dataset, using HRE (Home Range Extension – Rodgers and Carr, 1998). LSCV was used to automatically choose the smoothing parameter.

It is suggested that discretization be dealt with in the same manner as depicted in Figure 4.4.5. Since LSCV is based on calculated distances between points, the points themselves do not have to be moved, and the original dataset may be preserved (unlike the theoretical depiction in Figure 4.4.5.). Instead, if distances are zero (i.e. points are overlapping), then the distance may be manipulated artificially in the LSCV estimation (this process goes on behind the scenes). I propose that the rounding error be used as the basis for this manipulation, such that the rounding error forms the upper bound for a uniform distribution of numbers from which a manipulation distance is randomly selected. This is tantamount to shifting overlapping points within the area described by a radius equal to the rounding error. This does not allow for the opportunity of shifting points in the corners of grid cells, and this may in some cases be a problem. Sensitivity to this effect still needs to be evaluated.

I used an initial dataset of 50 random points, with x and y coordinates rounded to the nearest 10m (i.e. 5m rounding error), and then randomly selected 10 points to be repeated in subsequent analyses. Once the ratio of overlapping points to the total number of points reaches a certain threshold (Silverman, 1986), the smoothing factor (*h*) selected by LSCV will tend towards 0 (red data points, Figure 4.4.8.a.). This subsequently results in area estimates that also degenerate (Figure 4.4.8.b.). In both of these figures it can be seen that the manipulation of distances relative to the known rounding error will result in robust estimates of the smoothing parameter (blue data points).

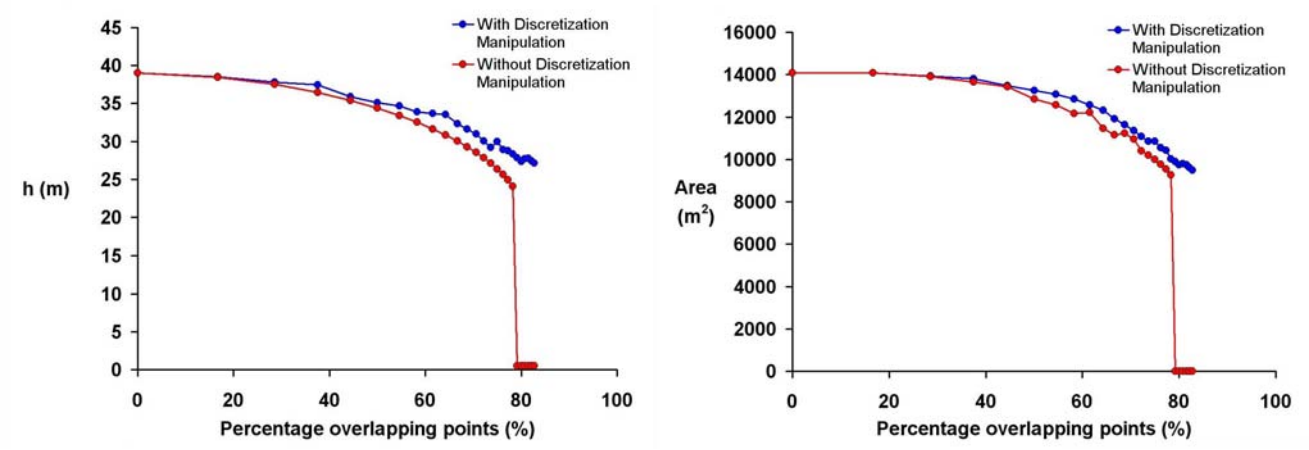


Figure 4.4.8. Degradation of *h* using LSCV with data that are discretized (10 separate locations show overlap from 1 to 8 points per location).

When only one point is repeated in the dataset, the breakdown of the LSCV estimate occurs much sooner (Figure 4.4.9.). It would seem from this cursory analysis that number of repitition ata single point may be more important than the number of points with repititions. Manipulation will suffice only up to a certain point, but ultimately, with enough discretization, the least-squares estimate will degenerate. It has been suggested that overlapping points should simply be removed from the dataset (Powell, Pers. Comm.). Where this treatment is not desired, perhaps a non-statistical method would be more appropriate. It is evident that **LSCV will not always be the most suitable method for choosing a smoothing parameter (especially with descretized data)**.

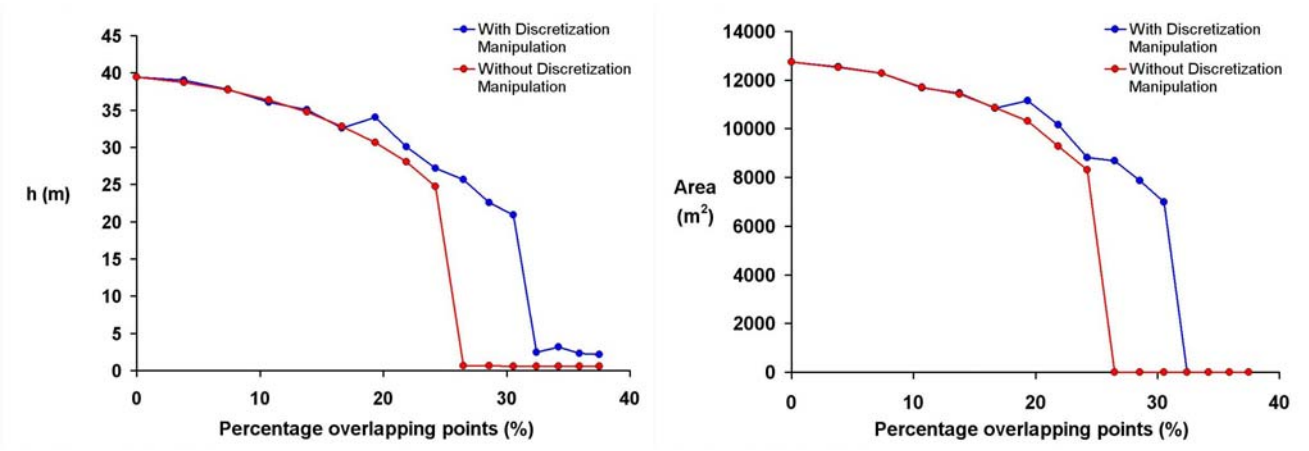


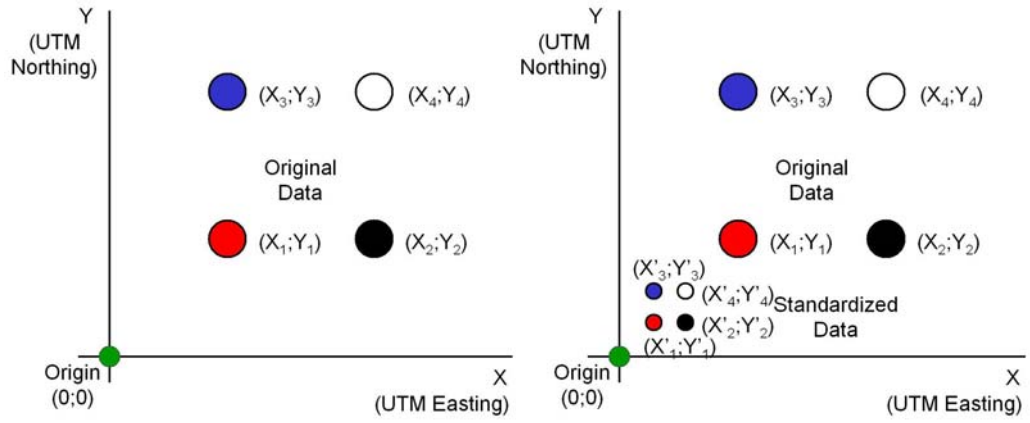
Figure 4.4.9. Degradation of *h* using LSCV with data that are discretized (1 location shows overlap from 1 to 38 points).

4.5. Standardization

Not all datasets are created equal. It is often the case that data will have greater variance in a particular direction. This may be caused by the behavior of the animal or it may be a result of a geographic constraint to movement – as might be the case for linear barriers such as mountain ranges, streams or coastlines (for terrestrial species). In such cases it may be better to search further in the direction with the greater variance to find points that will contribute to the density estimate for an evaluation point. Complex kernels are one method of overcoming this bias in the data, and an alternative technique is described in section 4.5.2. A better solution would be to standardize the data, run the kernel density analysis (using a single smoothing parameter) and then re-project the final home range contours to match the original scale of the data. Unit (Section 4.5.1.) and X Variance Standardization (Section 4.5.2.) have been suggested for this. ABODE allows the user to incorporate standardization in the analysis using the aforementioned methods and a thrid option, Covariance Bias (Section 4.5.3).

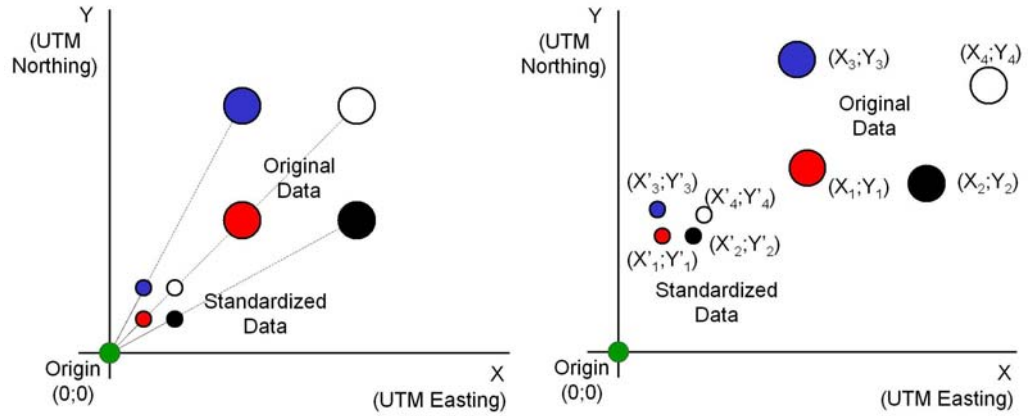
4.5.1. Unit Variance Standardization

Standardizing data to have a unit covariance matrix was proposed by Silverman (1986). The original data (Figure 4.5.1.1.a.) are standardized using the variance measures in the x and y directions. The x coordinate for each point is divided by the standard deviation in X (σ_x). Similarly, y is scaled by σ_y. This results in a set of standardized data (Figure 4.5.1.1.b.). In this simple example, the relationship between points is preserved, since the variance is equal in both X and Y (Figure 4.5.1.2.a.). In cases where the variance in X and Y is not equal (Figure 4.5.1.2.b.), the relationship between the points is altered such that the variance in each direction is equal. The kernel density estimation would typically be done on the standardized data to produce a home range estimate (Figure 4.5.1.3.a.). This allows for the use of a single smoothing factor. The final product would be transformed back to the original scale of the data (Figure 4.5.1.3.b.).



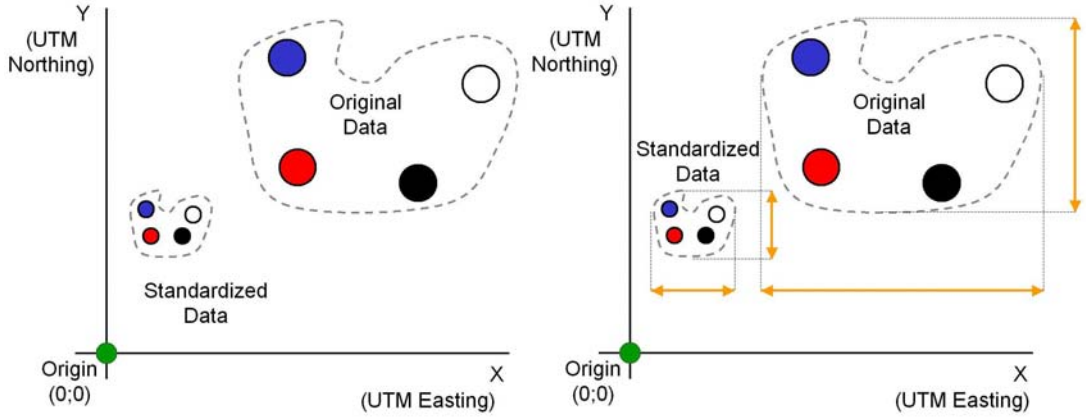
a b

Figure 4.5.1.1. Unit Variance Standardization – Original (a) and standardized datasets (b).



a b

Figure 4.5.1.2. Spatial relationships maintained with original datasets that have equal variance in the X and Y directions (a). Original data with unequal X and Y variances are transformed in a non-uniform fashion (b).

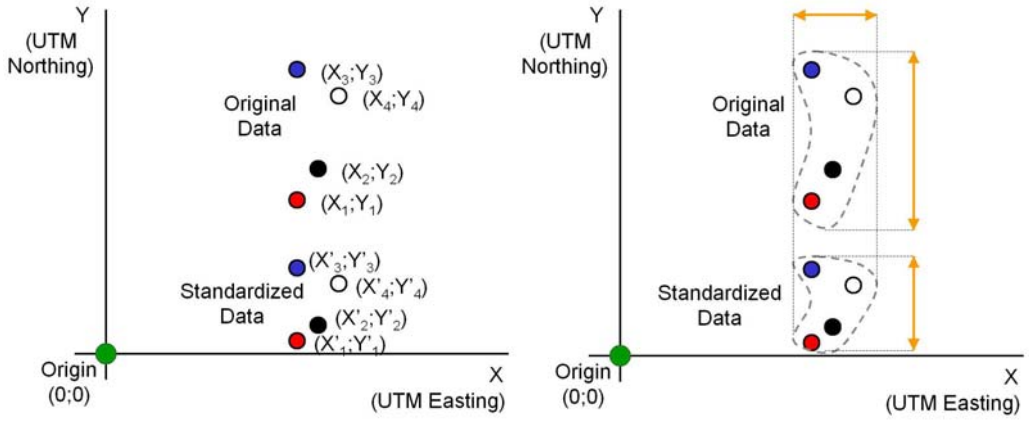


a b

Figure 4.5.1.3. Re-projection of home ranges (dotted lines) from standardized to original datasets using unit variance standardization.

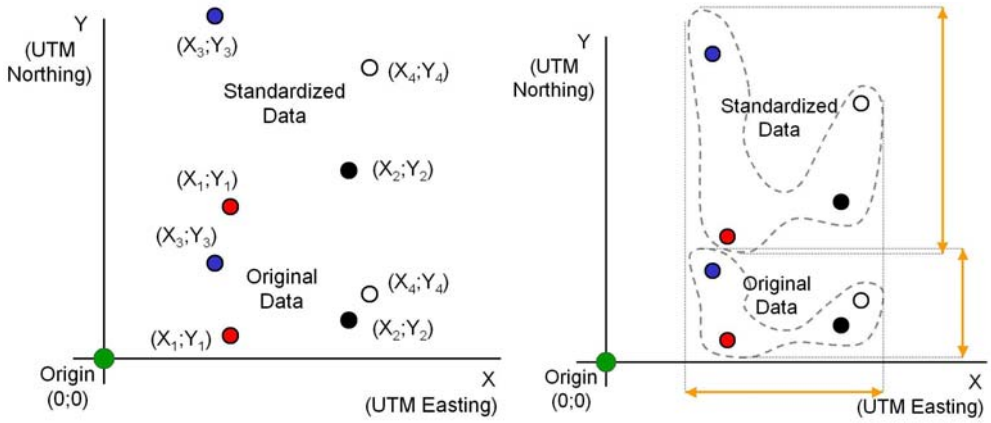
4.5.2. X Variance Standardization

Another method for equalizing the variance is to apply X Variance Standardization (Rodgers and Carr, 1998; Kenward and Hodder, 1996). The variance in the y direction is expressed in terms of the variance in the x direction. This is achieved by dividing the y coordinates by the standard deviation in Y and then multiplying them by the standard deviation in X (Figure 4.5.2.1.a.). Using this method, the variance in x is preserved. The home range obtained using the standardized data can then be re-projected using the factor of 0 for the x direction and σ_y for the y direction (Figure 4.5.2.1.b.). If the variance in y is greater than the variance in x, then the scale will be reduced in the standardized data (Figures 4.5.2.1.), whereas the scale will be expanded if the opposite is true (Figures 4.5.2.2.).



a b

Figure 4.5.2.1. X Variance standardization where the original variance in the y direction far exceeds that in the x direction.

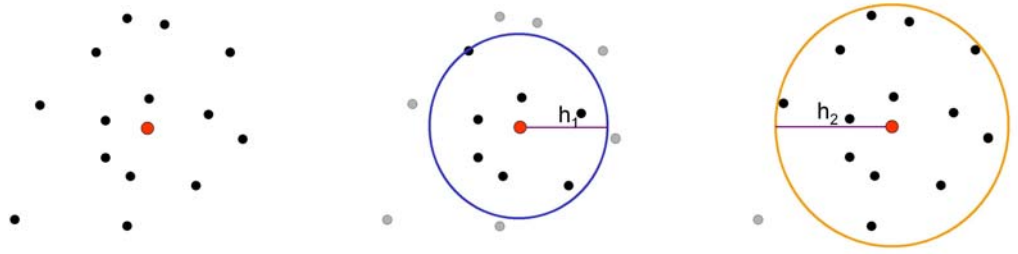


a b

Figure 4.5.2.2. X Variance standardization where the original variance in the x direction far exceeds that in the y direction.

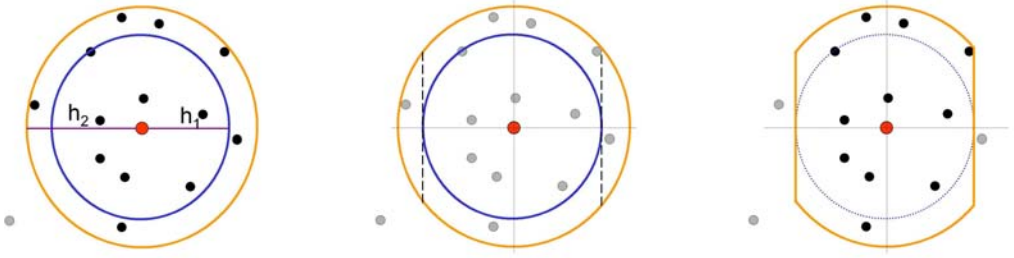
4.5.3. Covariance Bias

Seaman and Powell (1996) standardized their data to obtain an estimate of the required smoothing factor using the Least-Squares Cross-Validation technique. They then rescaled the smoothing factor to represent the variance in X and Y. This issue is dealt with in depth by Gitzen and Millspaugh (2003). ABODE provides an analog to the method used by Seaman and Powell (1996). Given a distribution of location estimates (Figure 4.5.3.1.a.), an evaluation point (red) is any point for which the kernel density is being estimated. The smoothing factor (h_1) (Figure 4.5.3.1.b.) represents the smaller of the variances (x in this case). Grey points would be excluded in the evaluation of the red point. Black points would be included. A second smoothing factor (h_2) represents the larger of the variances (y in this case) (Figure 4.5.3.1.c.). The single grey point would be excluded in the evaluation of the red point. Figure 4.5.3.2.a. shows the adjusted smoothing factors together. The actual search area is determined by a combination of the two smoothing factors (Figure 4.5.3.2.b.). The search area effectively becomes the circle of the larger smoothing factor (y in this case), truncated in the direction of the smaller variance (x in this case) (Figure 4.5.3.2.c.). Once the candidate points (black) have been obtained for the evaluation of the red point (Figure 4.5.3.3.a.), the larger of the smoothing factors is used in the weighting function applied (Equations 1 and 2). Distances of candidate points to the evaluation point are calculated (figure 4.5.3.3.b.). For the case of x variance greater than y variance the kernel would be oriented in a perpendicular fashion (Figure 4.5.3.3.c.).



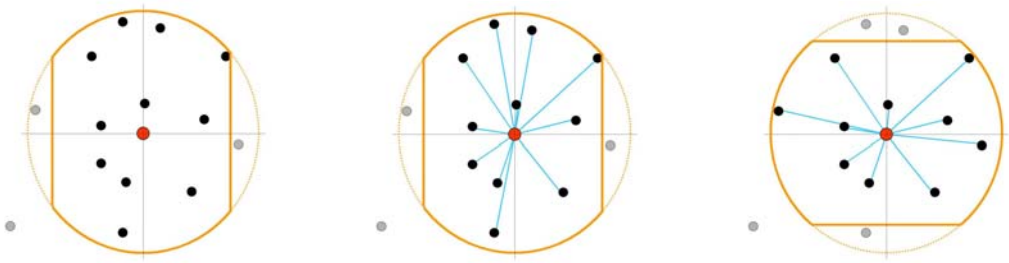
a b c

Figure 4.5.3.1. Generation of two smoothing parameters, based on either the x or y variance.



a b c

Figure 4.5.3.2. Truncation of the search area based on the smaller of the variances in either the x or y direction.



a b c

Figure 4.5.3.3. Inclusion of points in a density estimate based on the truncated search area when using the covariance bias method of standardization.

It should be noted that this method is not recommended. It is provided as an option in ABODE should the user wish to make comparisons between home range estimators that may only provide this method of standardization. This method is heavily biased by not incorporating the covariance between x and y.

5. Home Range Asymptotes

5.1. Why we should look at them

Harris *et al*. (1990) suggested that a home range analysis should be done using data that encompass the full range of variation in movement behavior attributable to sex and age differences. This is only possible if a representative sample (generally evenly spaced in time) is obtained for the entire sampling duration. To ensure that the sampling duration covers the full range of behavior exhibited by the animal, home range asymptotes are necessary. This should typically be done using a preliminary dataset, before the majority of the data are collected. The point in time (conversely the number of locations required) where the home range reaches an asymptote will indicate what the sampling duration (sample size) requirement should be. If home ranges do not asymptote, then the user may not have a representative sampling duration for the time period. Alternatively, a lack of asymptote may indicate a multiscaled home range (Gautestad and Mysterud, 1993). Harris *et al*. (1990) suggested the use of "area observation plots" (Otis and White, 1999) to determine the number of locations required to obtain a stable estimate of home range size (Stickel, 1954; Hawes, 1977). Gautestad and Mysterud (1995) proposed alternatively that home ranges are not asymptotic, but rather increase according to the power law (square root of the number of locations). This may be true for MCP analyses, but kernel estimators are relatively robust towards sample size issues (Seaman *et al*. 1999). Using simulation data, Seaman *et al*. (1999) showed that kernels gave stable estimates at about 50

locations for LSCV. They showed that kernel techniques actually overestimated home range size at lower sample sizes. They admitted that these results were obtained from simulated datasets, and it is unclear how real data will behave. Few analyses have used kernels in an asymptote analysis, mainly because of the tedium involved in doing them. ABODE provides an automated method for doing both MCP and kernel asymptotes.

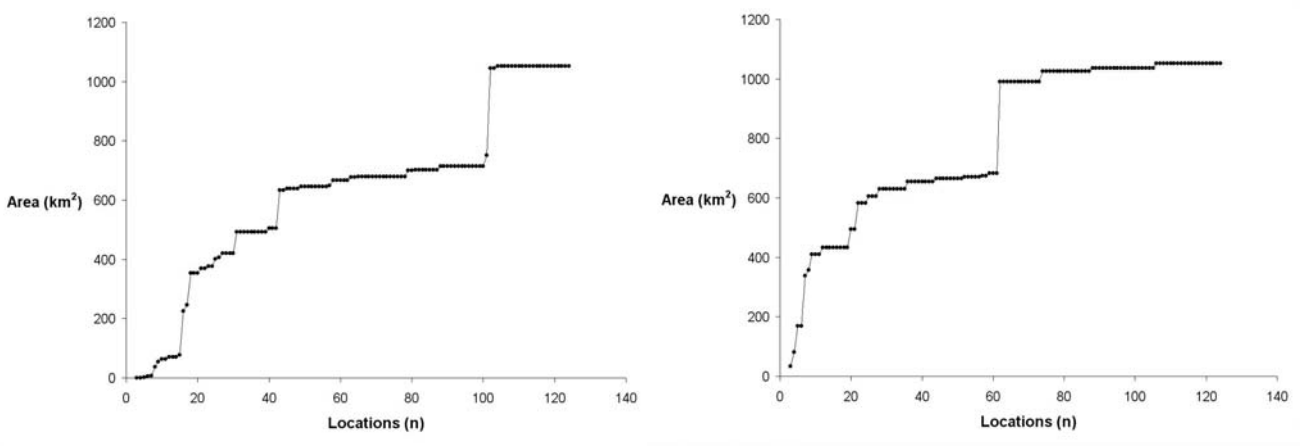
5.2. How we should analyze them

Home range asymptote analyses are carried out by estimating home range size at increasing sample sizes. This can be done by adding locations either randomly or sequentially (Harris *et al.*, 1990): if the data are continuous (i.e. they are collected at a constant sampling interval) then the locations should be added sequentially; if the data are discontinuous (irregular sampling interval) then the locations should be added randomly.

Forde (1989) used regression equations to correct non-asymptotic home ranges. If the sampling duration was appropriate then the fact that the home range does not reach an asymptote probably indicates a biologically important movement behavior. One home range that does not asymptote should not affect the estimate of required sample size. If many home ranges follow this pattern, then perhaps the behavior of the animal does not lend itself to a stabilization of home range size. The animal does not have what can be considered a true home range, but is instead migratory, continuously dispersing or simply transient.

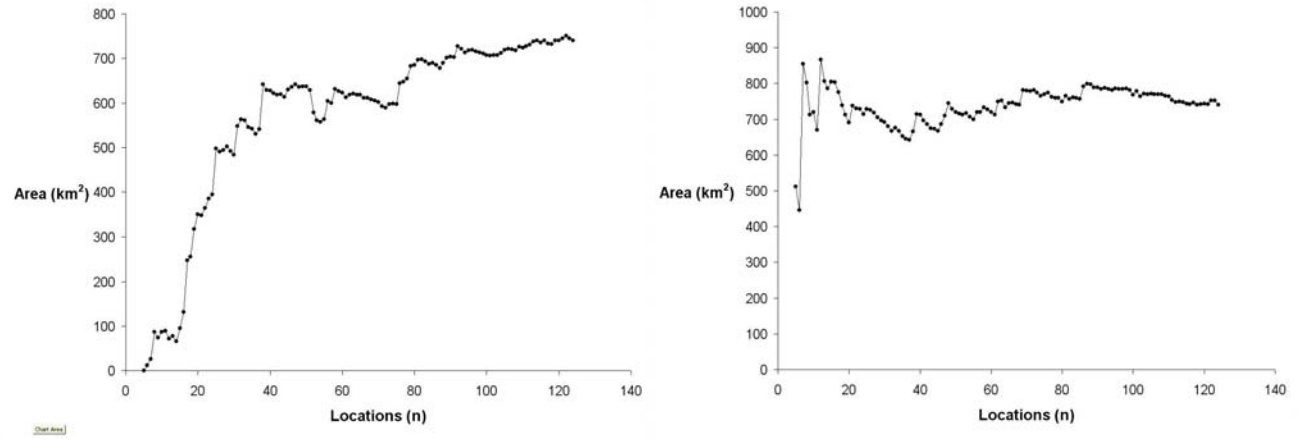
If the user chooses to do a kernel asymptote analysis, the option to output grids for each increment of locations will not be available. Outputting all the grids would take up too much space on most systems. The user will be able to output the shapefiles at the specified contour. For both kernel and MCP analyses, the most useful output from ABODE is the table that contains the area values for each subset of the data. These tables are labeled "_your shapefile name_" and then a suffix "KACons.dbf" (for the kernel asymptote using consecutive locations); "KARand.dbf" (for the kernel asymptote using randomly added locations); "MCPAcons.dbf" (for the MCP asymptote using consecutive locations); and "MCPArand.dbf" (for the MCP asymptote using randomly added locations). The next step in the analysis is to plot the area of each range estimate on the y-axis and the number of locations in the analysis on the x-axis. This can be done by opening the database file (.dbf) in a graphing package such as MSExcel. Graphically the user can then assess dispersal events or sallies (for consecutive MCP analyses) or the asymptote value and corresponding number of locations. No automated procedure is available for testing where a graph approaches the asymptote.

By using MCPs in the asymptote analysis, the user will be able to pick out dispersal events or significant sallies if the data are added sequentially (Figure 5.2.1.a.). A spike in the graph shows a significant increase in area. Finding the dates of such events and the associated age for each individual may be a way of estimating the dispersal age in a population. Estimating time gaps between spikes may indicate tenure times where territories are held. If the points are added randomly, this will not be possible (Figure 5.2.1.b.). Since kernel density estimates should be more robust towards outliers in the dataset, the detail required for analyzing individual behaviors is lost. The kernel asymptote analysis will however provide a better estimate for home range asymptotes using either consecutive locations (Figure 5.2.2.a.) or randomly added locations (Figure 5.2.2.b.).



a b

Figure 5.2.1. MCP asymptote analyses for one female cheetah using consecutively (a) and randomly added locations (b). Spikes in the graph indicate potential range shifts or exploratory sallies or perhaps dispersal.



a b

Figure 5.2.2. Kernel density asymptote analyses for one female cheetah using consecutively (a) and randomly added locations (b). Kernel estimation is relatively insensitive to range shifts, exploratory sallies or dispersal, and probably provides a better asymptote estimate if enough data are available.

6. Core Home Ranges

6.1. Does a core really exist?

Harris *et al*. (1990) concluded that core areas (if they exist in an animal's home range) may be useful in understanding the behavior of the animal, by providing a clearer interpretation of shifting patterns of use within a home range, and allowing better insight into intraspecific and interspecific patterns of area use. They suggested that in some cases total home ranges might overlap, while the core areas may be mutually exclusive. Powell (2000) described how core areas may indicate higher concentrations of important resources, and are thus more important to us in understanding an animal's life requisites than are peripheral areas. Not all animals will necessarily have a core in the range (Powell, 2000), and this could be due to an even or a random use of space by the animal. Cores should reflect biologically important areas in a range, rather than arbitrary probability cut-offs. In home range analyses, core areas are often reported simply as percentage use areas at some arbitrarily defined probability (Powell, 2000). This type of analysis should not be done. Rather, each home range should be tested to see if a biologically meaningful core area does exist.

6.2. How do we test for this?

Harris *et al*. (1990) proposed that core areas could be defined by plotting the area (y-axis) against the harmonic mean isopleth (x-axis) for harmonic mean analyses. The core area would be at the point of inflection on the graph. For kernel analyses, Powell (2000) and Horner and Powell (1990) suggested the use of a plot of the percentage area home range against the probability of use. The probability of use is plotted on the x-axis, scaled by the maximum probability of use (the highest probability of use occurs at 100%). The percentage area of the home range at a specified probability of use is plotted on the y-axis, scaled by the maximum area. In theory, random space use would result in a straight line plot (Figure 6.2.1.a.). 'Even' use of space would result in a convex curve (Figure 6.2.1.b.). Clumped use of space should result in a concave curve (6.2.1.c.). The core should be contoured at the probability of use where the curve has a slope equal to that for random use (m = -1) - this is equivalent to the point on the graph farthest (in vertical distance) from the slope m = -1.

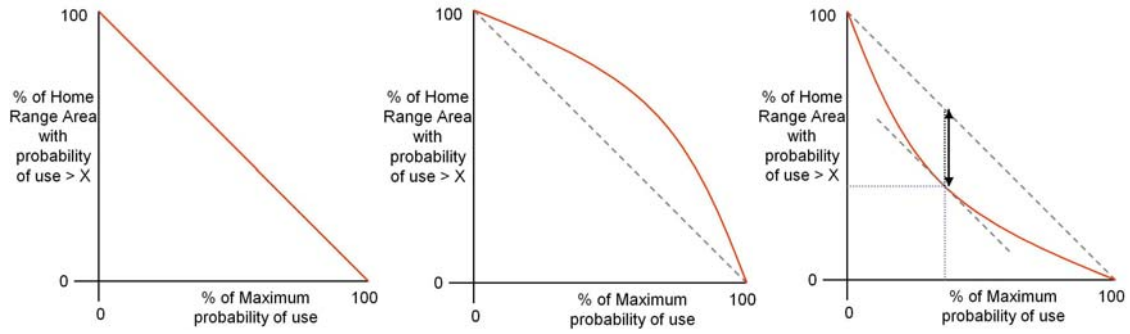
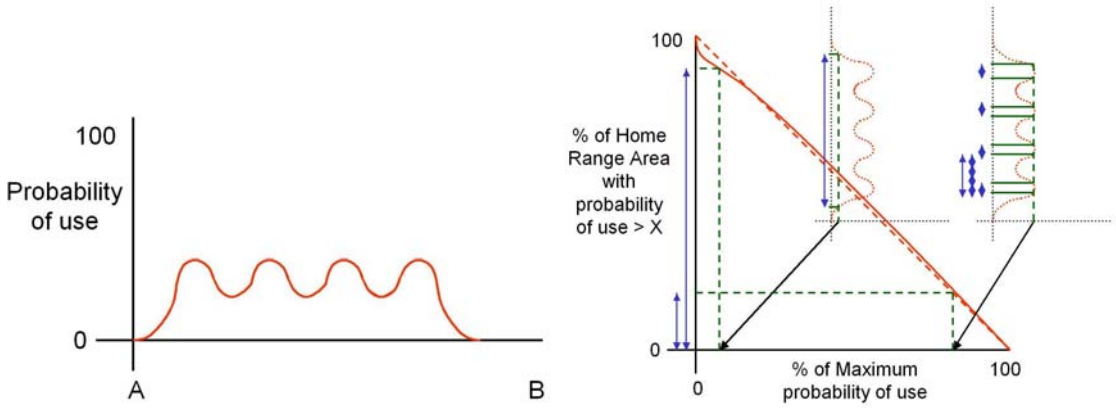


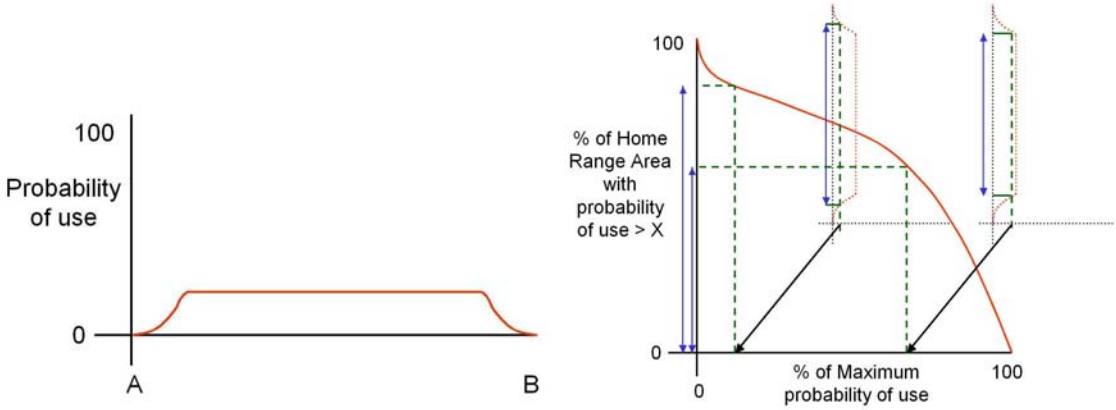
Figure 6.2.1. Core area determination following Powell (2000) and Horner and Powell (1990), with random (a), even (b) and clumped (c) use of space (Adapted from Powell, 2000).

In reality things are not this simple. Random use of space may result in small clumps in the data (Powell, 2000). At very high probabilities of use, there may be a concave shape in the graph. In addition to this, the distribution with random use, will not be everywhere random, since the location estimates fall-off at the periphery. So, at the edges of the distribution, the probability of use will be significantly lower than everywhere else within the middle of the distribution. Thus, a core area will be indicated for every distribution that has truly 'random' or 'even' use, only because the probability of use increases significantly just inside the edge of the distribution. Figure 6.2.2.a., is a cross section of a kernel density estimate for random use, where A and B are two points at the edge of the distribution. Plotted using Powell's (2000) method, the area contoured for a low probability of use will be large (high value on y-axis, but low value on the x-axis) while that contoured for a large probability of use will be much greater (Figure 6.2.2.b.). In Figure 6.2.2.b., the kernel density cross section is inverted and rotated so that the axes of the probability plot and that of the kernel plot will roughly match up. In theory, an even distribution should produce a graph that is flat (equal area value) for all probability of use values. There should only be one value for probability of use, since it is everywhere equal. In reality, at the edge of the distribution probability of use will be lower than the middle but will increase slightly before remaining constant, moving inwards from the edge (Figure 6.2.3.a.). The result of this kernel density cross section can be seen in Figure 6.2.3.b., where there will be a core estimated at a very low probability of use (near the boundary of the used area). Finally, in a clumped distribution, as with normally distributed locations (Figure 6.2.4.a.), a true concave pattern is evident in the probability plot (Figure 6.2.4.b.). Here the point where the core should be contoured is where the probability of use increases significantly (indicated by a slope m = -1).



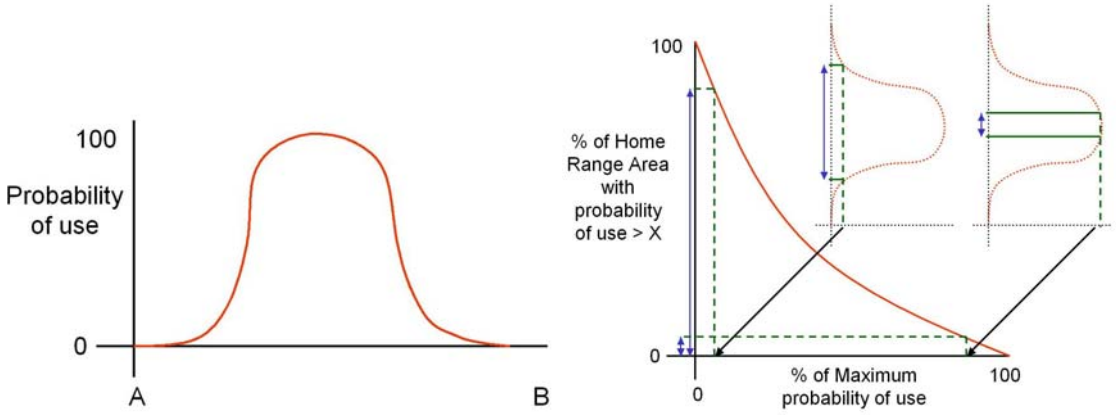
a b

Figure 6.2.2. Methodology for generating a core area probability plot with a random distribution of locations.



a b

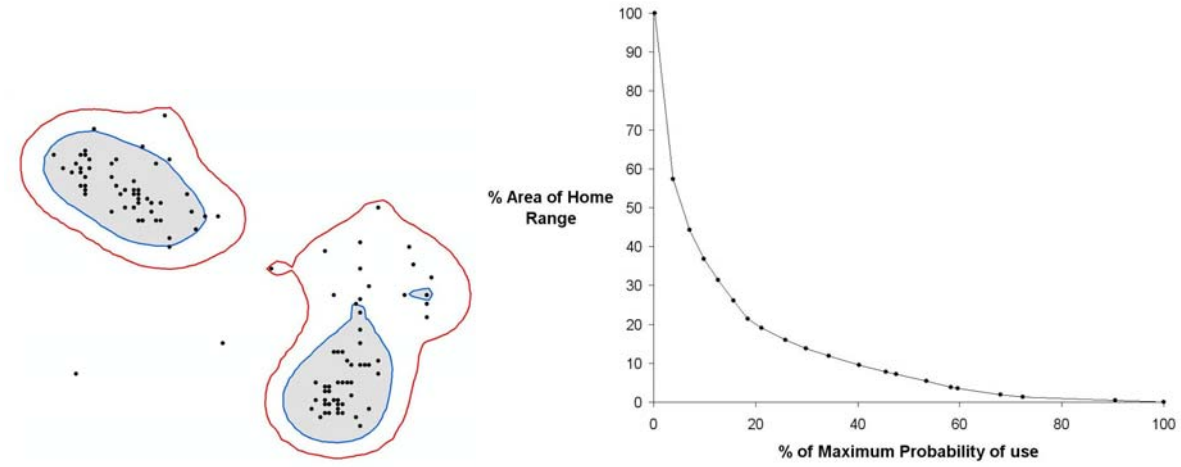
Figure 6.2.3. Core area probability plot with an even distribution of locations.



a b

Figure 6.2.4. Core area probability plot with a clumped distribution of locations.

It is evident that all distributions will exhibit home range cores using this method. Random and even distributions of use will show core areas (according to the definition) even if these appear within but near the periphery of the distribution. In such cases, this could be used as the definition of the total home range (instead of the arbitrarily defined 95% volume cut-off). For random distributions, cores may be exhibited for very high probabilities of use, and very small areas. These cores are probably not biologically meaningful. In all cases it is suggested that the user plot these graphs for every core area analysis. A decision based on the general shape of the graph should guide the interpretation of core areas. In other words, where the graph appears to be generally straight with a slope of m = -1 until very high probability of use, then the animal is probably using space randomly. In ABODE, core areas will be contoured based on the above method. This means that a shapefile of the core (whether artifact or real) will always be produced (Figure 6.2.5.a.). The user is advised to then plot the table that contains the areas and probabilities. It is called "_your shapefile name_" and a suffix of "CoreGraph.dbf", and can be opened in a graphing package such as MSExcel (Figure 6.2.5.b.). This *post hoc* visualization is recommended for every analysis. It should be noted that the x-axis in such a plot is not the percentage home range, but rather the percentage of the greatest probability value in the density grid.



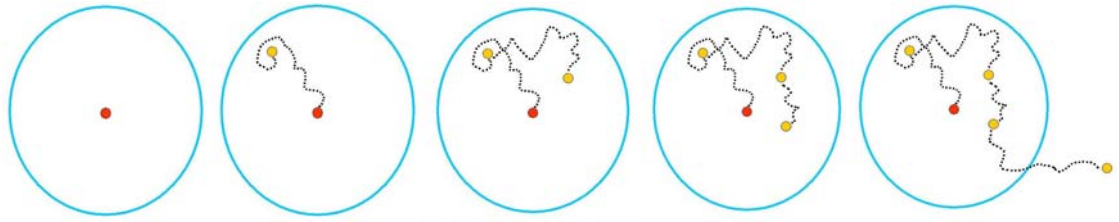
a b

Figure 6.2.5. 95% kernel home range (red line) and core home range (contoured in blue at 70% of the volume of the density surface) for a single female cheetah. Both estimates used LSCV, fixed kernels, unit variance standardization, and 1000m grids. The core area probability plot used to determine the core area shows a concave form.

7. Data driven and Biologically meaningful methods

It has been shown that LSCV will not always be a suitable method for selecting a smoothing parameter automatically (Silverman, 1986). It is also inappropriate to use a subjectively chosen smoothing factor when comparison of home ranges is desired. In many cases, our data and the

behavior of our study species will give us clues as to what our smoothing parameter should be. When obtaining the kernel density estimate, the smoothing parameter (h) determines which points will contribute to the density at any particular point in space (since it defines the search radius around each point). It makes intuitive sense that areas that are biologically important should have a high density of use. If a point is biologically important, then the animal will keep using and/or coming back to that place – this results in the higher density. If we artificially choose a smoothing parameter that includes too much area (large value for h), then an artificially high density may be accrued because the algorithm searched too far to find points that would contribute to density. The distribution of data for an animal's home range will usually be determined by a combination of the sampling interval (and sampling duration), and the behavior or natural history of the animal. Some data are collected on a regular sampling schedule, with a generally constant sampling interval. Given this type of schedule, the next location estimate for the animal will be determined by the previous location estimate and by the average distance that the animal would move within the given sampling interval. One would expect that the animal would be found somewhere within this radius. Some other behavioral phenomenon may determine the shape of the search area, which might not be a circle. If a smoothing parameter is chosen that far exceeds this hypothesized radius (Figure 7.1.a.), then subsequent location estimates for that animal may be included in the density estimate of a point even though the animal was only passing through the area. In other words, location estimates at subsequent sampling efforts are included only because the animal would not have had enough time to move out of the area before being sampled again. This is depicted in Figures 7.1.b-e, where an animal's walk (dotted line) is traced after having left the evaluation point (red). The subsequent location estimates (yellow) are the places where the animal was sampled at a fixed interval. The density estimate for this point (red) would then be artificially high since the place evaluated may not have been biologically important to the animal, but it was sampled often there because of the relationship between its movement behavior and the sampling interval.

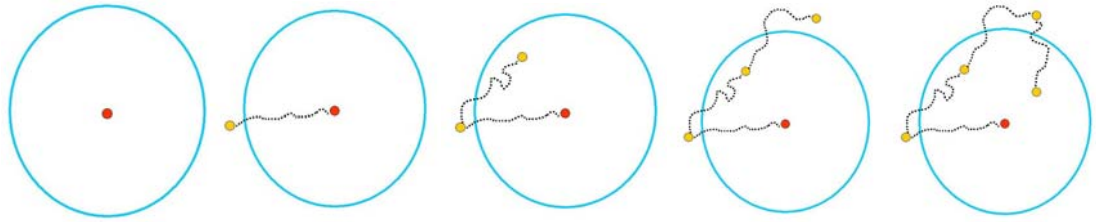


a b c d e

Figure 7.1. Inappropriate choice of smoothing parameter allowing the inclusion of points from subsequent days as an artifact of the daily distance traveled.

A better method for analysis would be to choose a smoothing parameter that reflects the relationship between the animal's movement behavior and the sampling interval. If *h* is chosen

equal to or smaller than the distance typically traveled in the given sampling interval (displacement distance)(Figure 7.2.a.), then in theory, the animal would have left the area by the next sampling event (Figure 7.2.b.). Resightings in that area should only occur if the animal chose to stay in the area instead of moving its typical interval distance, or if it subsequently returned to the area (Figures 7.2.c-e). In both cases it is fair to assume that a high density of location estimates may reflect biological importance to the animal.



a b c d e

Figure 7.2. Appropriate choice of smoothing parameter based on the structure of the location data (sampling schedule) and the movement behavior of the species (displacement distance).

ABODE provides a smoothing selection function in which the user can enter the sampling interval for the given dataset. This is the smoothing function "Displacement". If ancillary data that include the day, month, and year of each location are provided, then ABODE will search the dataset and find all consecutive location estimates for the given interval. The distance traveled between the locations will be calculated, and the arithmetic mean thereof will be used as the smoothing parameter for the dataset. In reality, not all studies will have the luxury of a generally consistent sampling interval. Many data are collected opportunistically, and will thus have an irregular sampling schedule. In such cases, ABODE can look for a sampling interval that best represents the data. The user should decide which measure of central tendency would best represent the data. ABODE can look for the median, mean and mode sampling interval. Before making a selection, the user should understand the implications that outliers in the data will have on each estimate. It may be best to run all three analyses and choose a method that seems most appropriate.

Sample size considerations dictate that ABODE uses a range of values for its search criteria. The mean sampling interval may not even be represented in the dataset, or may, at best, be represented by only a few actual data pairs. To base a travel distance on such a small sample size may produce undesirable results. Similarly, the median sampling interval may only be represented once in a dataset, and the mode, at least twice. ABODE will search for sampling intervals within a range that operates on an increasing scale (Table 1.). If the interval ABODE is searching for is less than 7 days, then the tolerance (on either side) is 0 days. If the interval is

less than two weeks (14 days) then the tolerance is 1 day on either side. This means that if the interval is one week, ABODE will search for data pairs that are 6, 7, or 8 days apart, and then calculate the displacement from the new set. These values were chosen subjectively by the author and could be adjusted very easily in the code for the program.

Table 1. Tolerances used to guide search criteria in the "Displacement" Smoothing Function option in ABODE. The interval that the search is based on is given in days, and the tolerance is the number days on either side of the given interval that ABODE will use when searching for displacement data pairs.

Interval (Days)	Tolerance (+/-Days)
<7	0
7 - 13	1
13 - 29	2
30 - 59	4
60 - 89	8
90 - 119	8
120 - 149	15
150 - 179	15
180 - 359	15
>360	30

8. Using ABODE

ABODE is a Home Range tool developed for ArcGIS using ArcObjects and Visual Basic for Applications (VBA). To use ABODE, you will need ArcGIS 8x or higher. The code and VBA form provided will not work in ArcView. The functionality of ABODE is limited by the user's understanding of home range analysis –**therefore it is suggested that the first part of this user manual be understood**. The use of ABODE also requires that users have a basic understanding of and proficiency in ArcGIS (much of the terminology relating to this software package has been left undefined in this user manual). Default settings were purposefully omitted from the program in an attempt to force users to make well-informed decisions about their analyses. This user manual is designed to give a very basic understanding of home range analysis such that a user will be able to decide upon the best inputs and constraints for the analysis.

ABODE can be used as a Visual Basic for Applications (VBA) user form, that can be added to an existing map document (mxd) or it can be used in a map document or map template (mxt) to which the form has already been added. Both formats are provided and explained in the following sections. The easiest method would be to use the map document supplied, and users wishing to do so should skip Section 8.1.1. and Section 8.1.2.

8.1. How to start using ABODE

8.1.1. Loading the form into the VBEditor

You have two choices in applying this program. You can either save ABODE as part of a template document or you can save ABODE independently as part of a unique and saved map document (.mxd). Be warned that if you save ABODE as a template, it will appear as an available tool every time you open a new empty map.

Right click on the menu bar (Figure 8.1.1.1.). Scroll down to the "Customize…" option and left click on it (Figure 8.1.1.2.). Left click on the "Commands" tab, scroll under "Categories" and left click on "UIControls" (Figure 8.1.1.3.).

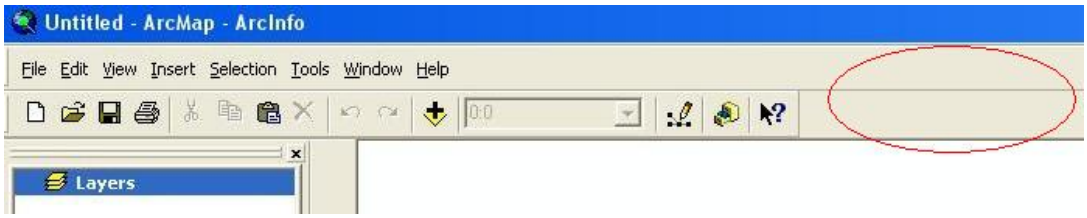


Figure 8.1.1.1.

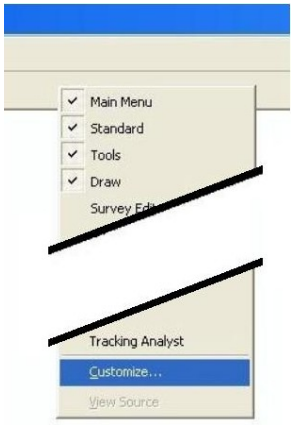


Figure 8.1.1.2.

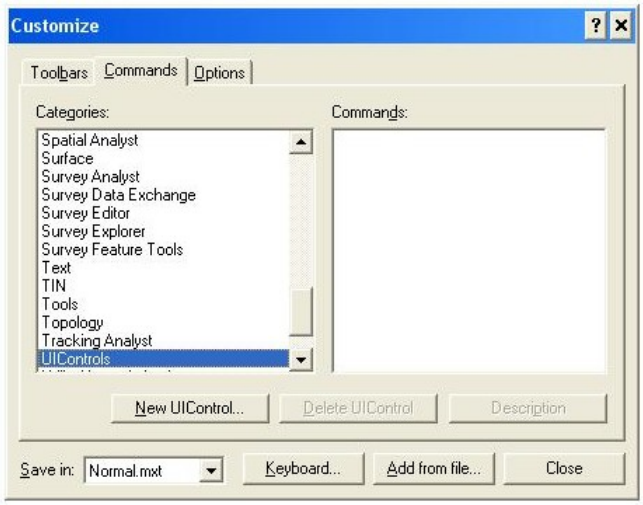


Figure 8.1.1.3.

You will notice that the "Save in:" option reads "Normal.mxt" if you have not saved the map document yet (Figure 8.1.1.4.). Use this option if you wish to have ABODE appear every time you open a new and empty map. This is the "normal" template to which the code and form will be saved and subsequently applied. Without having saved the map document yet, the other option

in the drop-down menu reads “Untitled” (Figure 8.1.1.5.). If you have saved your document, the option will show your document’s name (Figure 8.1.1.6.). Left click “New UIControl…”. From the options given, leave the radio button for “UIButtonControl” activated (Figure 8.1.1.7.).

Figure 8.1.1.4.

Figure 8.1.1.5.

Figure 8.1.1.6.

Figure 8.1.1.7.

Left click “Create”. In the “Commands” window, “Normal.UIButtonControl1” is displayed if you have chosen the template option; “Project.UIButtonControl1” is displayed if you have chosen to save as part of a unique document. Change the extension “.UIButtonControl” to “.ABODE” (Figure 8.1.1.8.). Drag the newly created button to a tool bar (Figure 8.1.1.9.). Right click on the new button. Check “Text Only” (Figure 8.1.1.10.). This should display the tool as the name entered for the extension.

Figure 8.1.1.8.

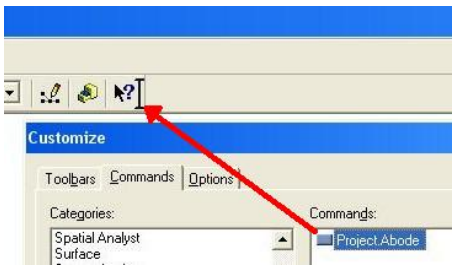


Figure 8.1.1.9.

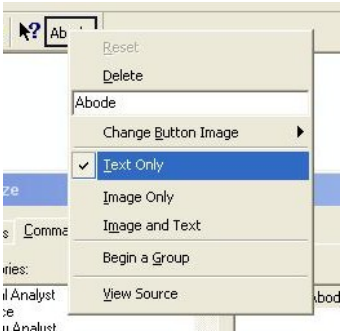


Figure 8.1.1.10.

8.1.2. The VBA realm

With the “Customize” window open, you can right click on the tool (ABODE button) and choose the “View Source” option to enter the Visual Basic Editor (Figure 8.1.2.1.). Alternatively, with the “Customize” window closed, you can right click on the tool (whenever the ABODE window is closed) and scroll down to choose the “View Source” option (Figure 8.1.2.2.). A left click on “View Source” will open the Visual Basic Editor (Figure 8.1.2.3.).

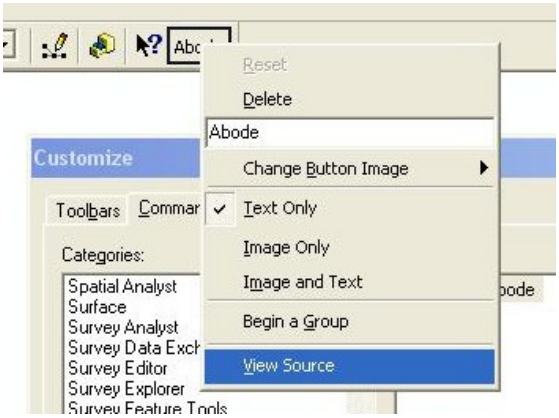


Figure 8.1.2.1.

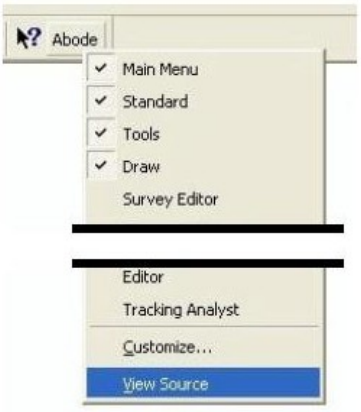


Figure 8.1.2.2.

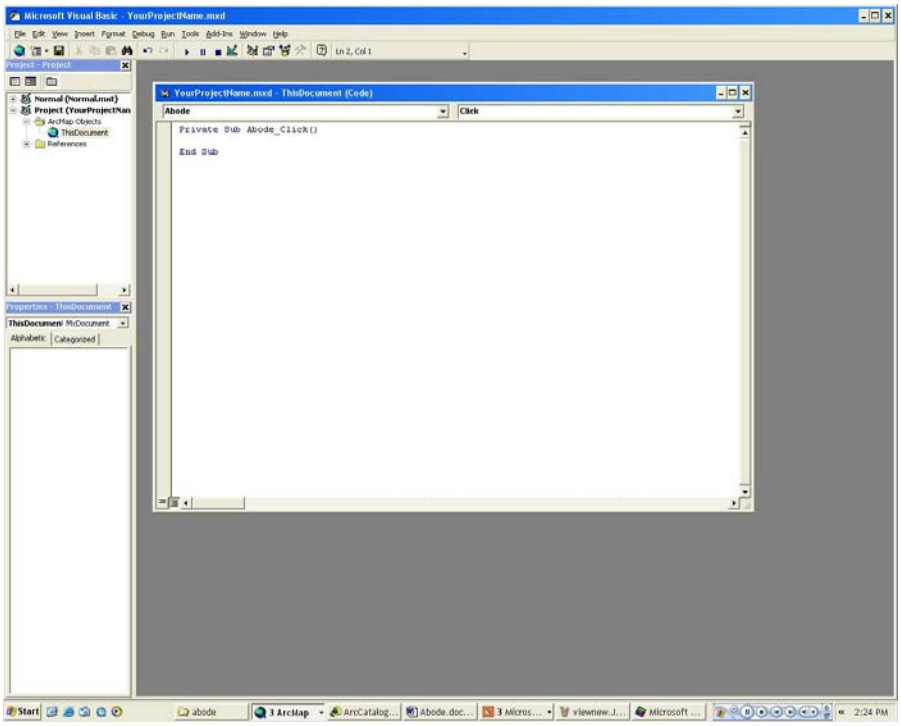


Figure 8.1.2.3.

The active window is the “ThisDocument” code source. It will read “Normal.mxt” or the name of your document depending on the option chose above (Figure 8.1.2.4.). Without clicking anywhere, you may begin typing to have the code inserted between the ‘wrappers’. Wrappers contain the code for calling the form you are about add, whenever the tool button is left clicked. Between “Private Sub _your tool name_click( )” and “End Sub”, where the cursor currently rests, type the following (Figure 8.1.2.5.):

load frmhr

frmhr.show

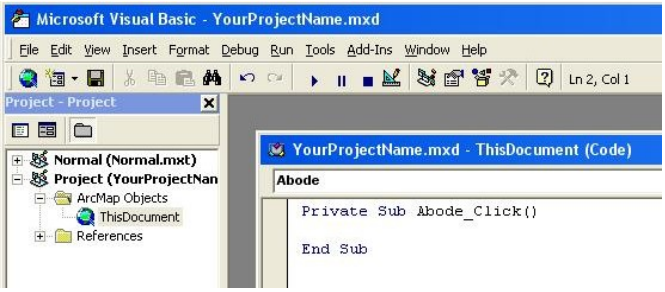


Figure 8.1.2.4.

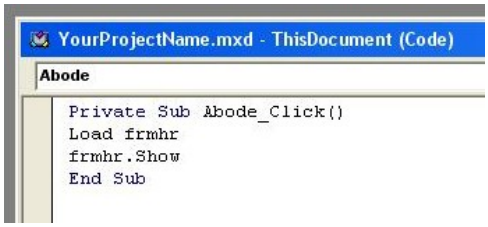


Figure 8.1.2.5.

Now right click in the Project explorer window, and left click on the “Import File…” option (Figure 8.1.2.6.). Navigate to where you saved the form (ABODE.frm) and open this file (Figure 8.1.2.7.). Expand the folder for “Forms”, and you will see that “frmHR” is registered (Figure 8.1.2.8.).

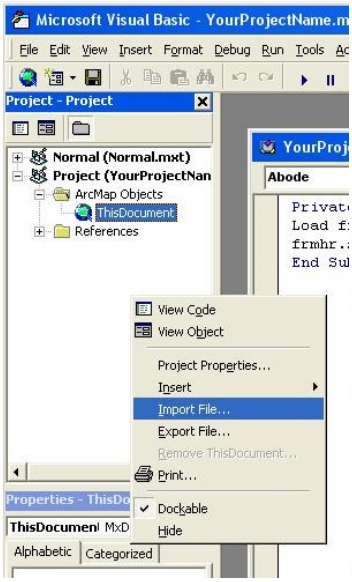


Figure 8.1.2.6.

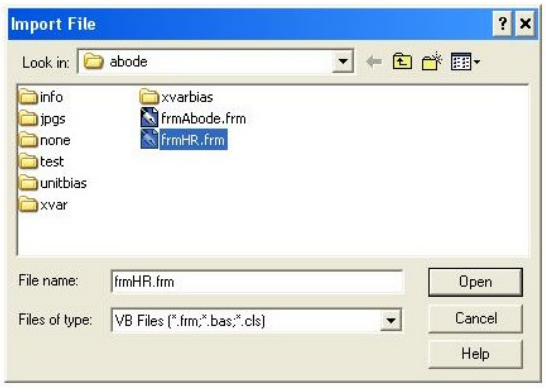


Figure 8.1.2.7.

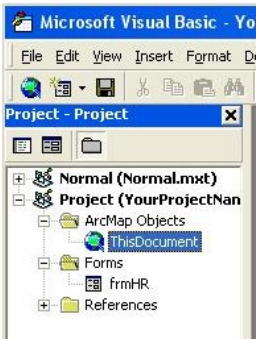


Figure 8.1.2.8.

If you double left click on this option, the form that runs ABODE will be displayed (Figure 8.1.2.9.). If you then go back to the “ThisDocument” code, with a double click, you will see that capitalization of the form name has occurred and it will now read as

Load frmHR

frmHR.show

indicating that the form has been recognized (Figure 8.1.2.10.).

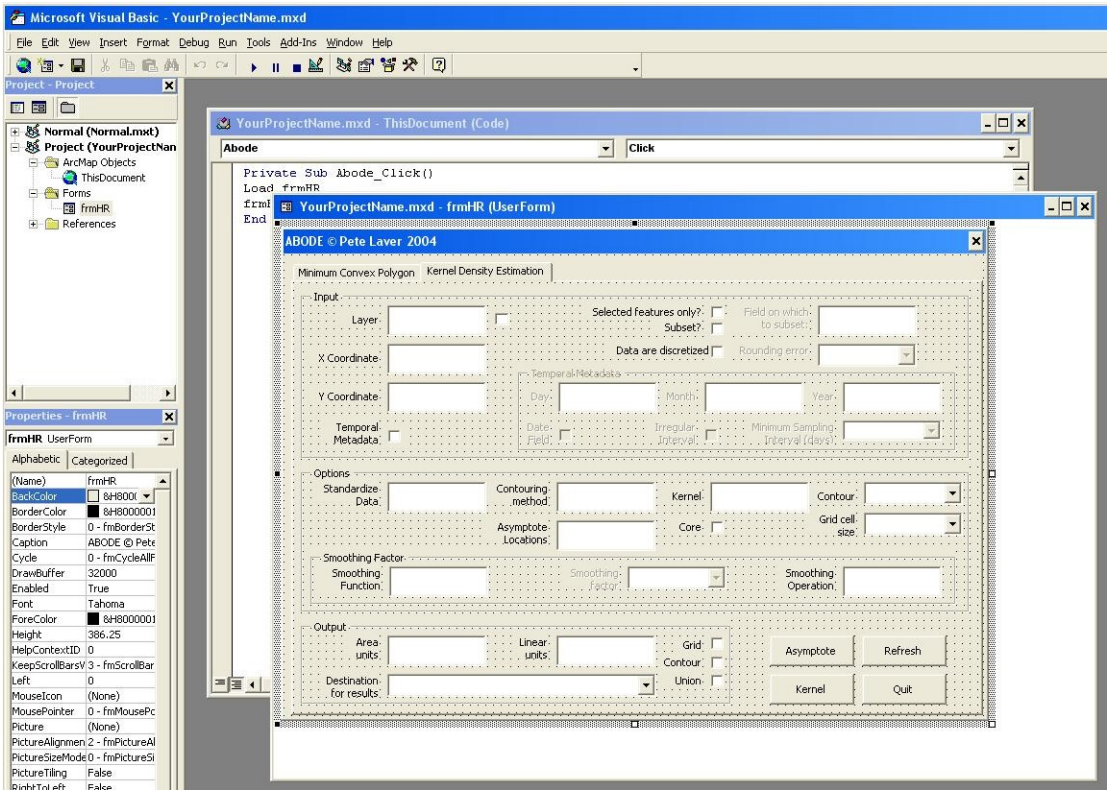


Figure 8.1.2.9.

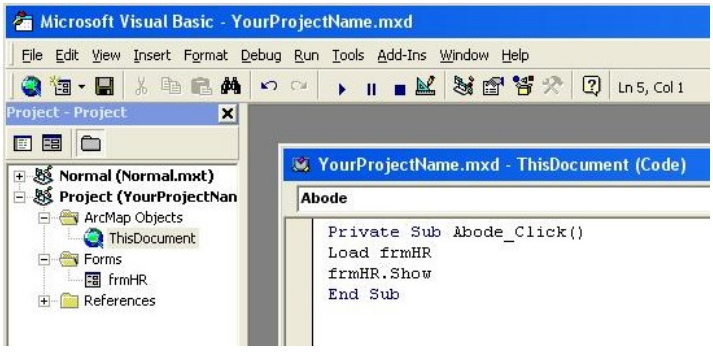


Figure 8.1.2.10.

One more step is required to start ABODE. On the menu bar, left click on “Tools”. Left click on “References…” (Figure 8.1.2.11.).

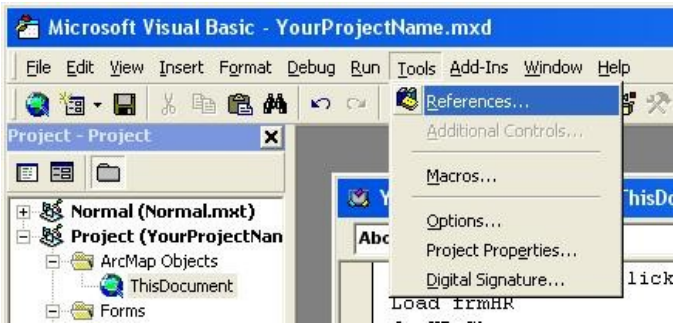


Figure 8.1.2.11.

Scroll through “Available References”. If you are using ArcMAP 8x, Check “ESRI Spatial Analyst Extension Object Library” and “ESRI Spatial Analyst Shared Object Library” (Figure 8.1.2.12.). If you are using ArcMAP 9, these references will be denoted as “missing” and they should be unchecked. Close the Visual Basic Editor.

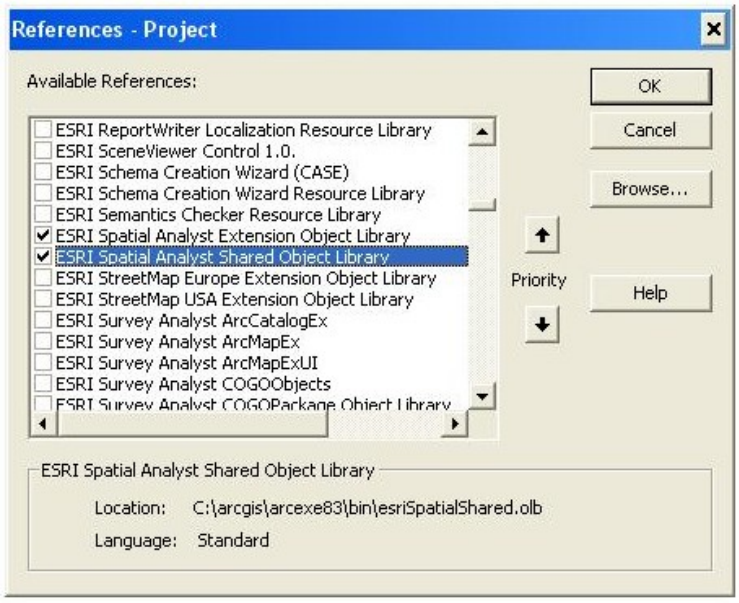


Figure 8.1.2.12. Referencing object libraries in ArcMAP 8x.

8.1.3. The easy start-up for ABODE

ABODE has already been loaded into a map document and a map template called “abode.mxd” and “abode.mxt” respectively. Referencing of Object Libraries still has to be performed. Figure 8.1.2.11 and its supporting text in Section 8.1.2. explain how this can be done. The map

document or map template may be saved with other names. Data should be loaded into the document in the form of a shapefile. It is recommended that the data are projected in UTM meters (ABODE has not been tested with other projections systems yet).

8.2. Using ABODE for home range analysis

In order to facilitate the automation of tedious home range analyses, ABODE provides options for subsetting datasets. This is equivalent to running batch mode. This allows users to load their data as a single shapefile into ArcMAP. As an example, one shapefile may have the location data (x and y coordinates), day, month, year, and ID for all the animals in a population. When subsetting, the user chooses the field on which to subset (i.e. the ID field), and ABODE will then run the anaylsis for the data that have the same value in that field, before starting new analysis for the next value. Thus, if a dataset has 5 individuals, with IDs "1", "2", "3", "4" and "5", and if the field "ID" is chosen for subsetting, ABODE will run the analysis first for the records for individual "1" then for individual "2" etc, and then provide a single table as output containing the results from the 5 analyses. ABODE can only handle subsetting at one level, but if separate analyses are desired for multiple fields, then the fields in the dataset should be manipulated in either ArcMAP in the attribute table or in a software package such as MS Excel, MS Access or SAS, before being imported back into ArcMAP. In this manner, a yearly analysis by individual is possible by generated a field (i.e. "IDyear") in which the ID and year are concatenated (individual "1" in year 83 would be 183, and year 84 would be 184 etc). Since some of the options in ABODE are date sensitive (they require day, month and year fields) it is suggested that a date field (if available) be converted to separate fields for day, month and year. This can be done in MS Excel using the functions '=day()', '=month()' and '=year()'.

8.2.1. The Visual Basic form and error trapping

ABODE was designed to be a relatively simple program with open code. The benefit of this feature is customization. The field of home range analysis is progressing fairly rapidly. No estimation techniques are or should be taken as irrefutable truths. Thus, it is anticipated that the technology and theory behind home range estimation will change. In addition to this, no technique will fit all data. Instead of 'massaging' data to fit a technique, techniques should be developed and changed such that we are faithful to our data and their capabilities. Written in Visual Basic for Applications and ArcObjects, you as a user have the ability to change the code. The user interface is a simple form in which you enter or select parameters to be used in the analysis.

ABODE has been tested, but as with most software, it has not been tested enough. The form is designed in such a way that it traps potential errors before the code runs. This prevents the user having to debug code line by line – a tedious and difficult procedure. The drawback is that the form may be a source of frustration to many users who are trying to customize their analysis. In most cases the form will not allow procedures that may cause the program to crash. This often occurs when the user has not chosen a required input. When the form fails to catch these potential errors however, and error message will occur, that asks the user if they wish to 'debug' (Figure 8.2.1.1.), if the "Allow Debug" checkbox is checked.

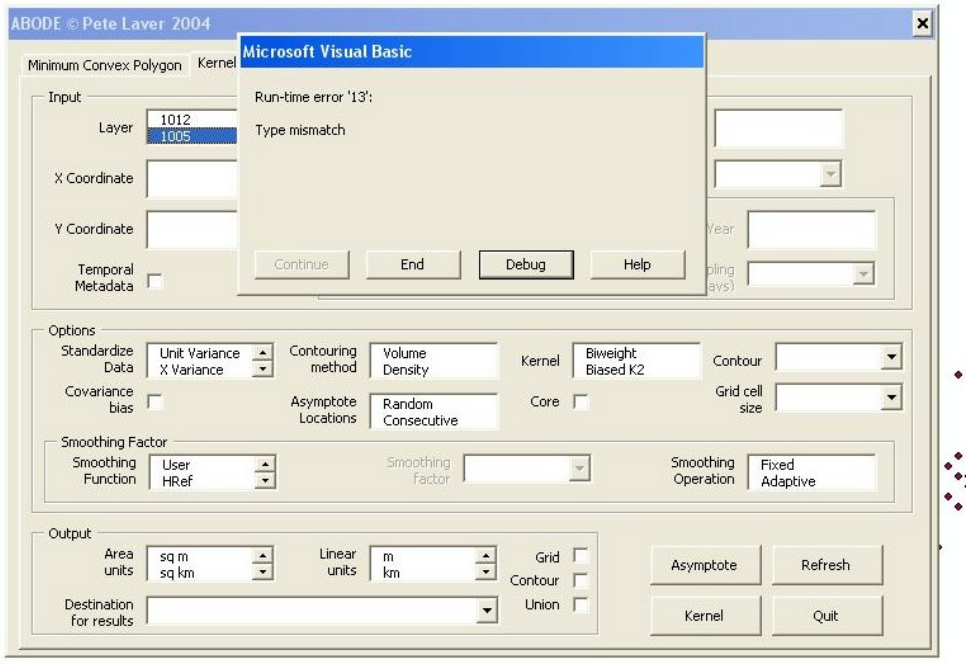
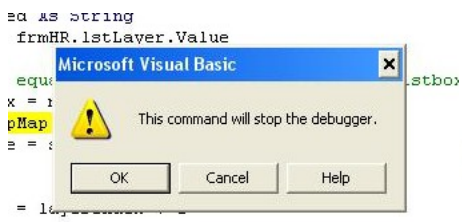


Figure 8.2.1.1.

The user may choose to debug (this is not recommended) or may choose 'end' (recommended), and determine based on the attempted inputs why the procedure was not allowed. Choosing to debug will open the Visual Basic Editor. If the user wishes to exit at this stage they may close the editor. At this point they will be prompted with stopping the debugger (Figure 8.2.1.2.). Choosing "Cancel" will return the user to the visual Basic Editor. Choosing "OK" will close the editor and return the user to ArcMap, closing the failed ABODE procedure. If you forgot to check the "ESRI Spatial Analyst Extension Object Library" and "ESRI Spatial Analyst Shared Object Library" (Figure 8.1.2.12.), you will get an error when you first run a kernel analysis (Figure 8.2.1.3.). In this situation, click "OK". Go through the steps outlined above (Section 8.1.2.) for setting the required references. Then close the Visual Basic Editor. Go to the folder you chose for saving

the outputs from the analysis. Delete any new table that has been created. Begin your analysis again.



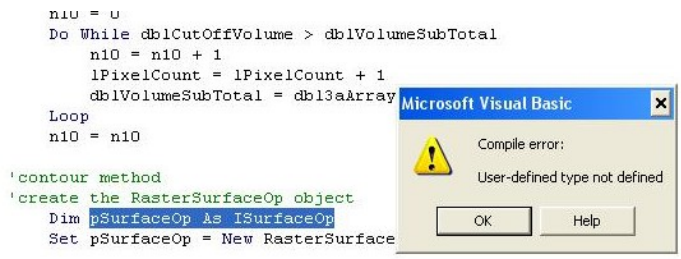


Figure 8.2.1.2.

Figure 8.2.1.3.

If the “Show errors” checkbox is left checked, a message box will pop up on every error with a description. This may be left unchecked if multiple analyses are performed using the Subset option – i.e. the code will not be interrupted, and the analysis will run to completion without notifying the user of the errors.

8.2.2. Minimum Convex Polygons (MCPs)

To begin a home range analysis, a point layer must be present in the Table of Contents in ArcMap. Start the analysis by left clicking on the toolbar, on the tool called “ABODE”. If no point layer is present in the Table of Contents, then ABODE will not open the user form. Instead a Message Box will appear notifying you that a point layer must be added. ABODE can analyze point shapefiles that have points projected in UTM. In the attribute table of the shapefile, there must be a field each for the x and y coordinates.

The user form for ABODE consists of two pages. One is for Minimum Convex Polygons (Figure 8.2.2.1.). The other is for Kernel Density Estimation (this will be dealt with in section 8.2.3.). This page is divided into two distinct sections. The top section is for regular Minimum Convex Polygon analyses (Figure 8.2.2.2.). The bottom section is for analyzing home range asymptotes (Please note that on the Kernel Density Estimation page, all analyses are included in one section, and asymptotic analyses are not set apart).

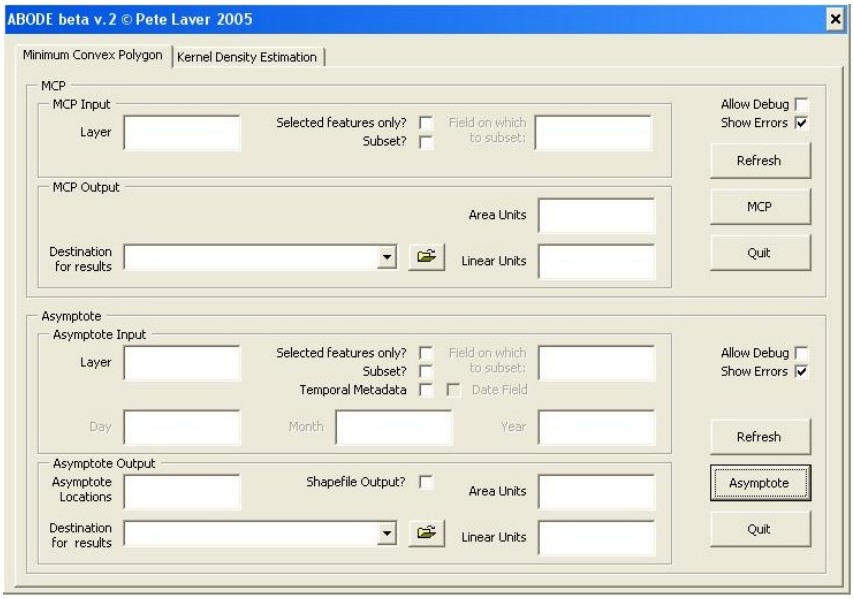


Figure 8.2.2.1.

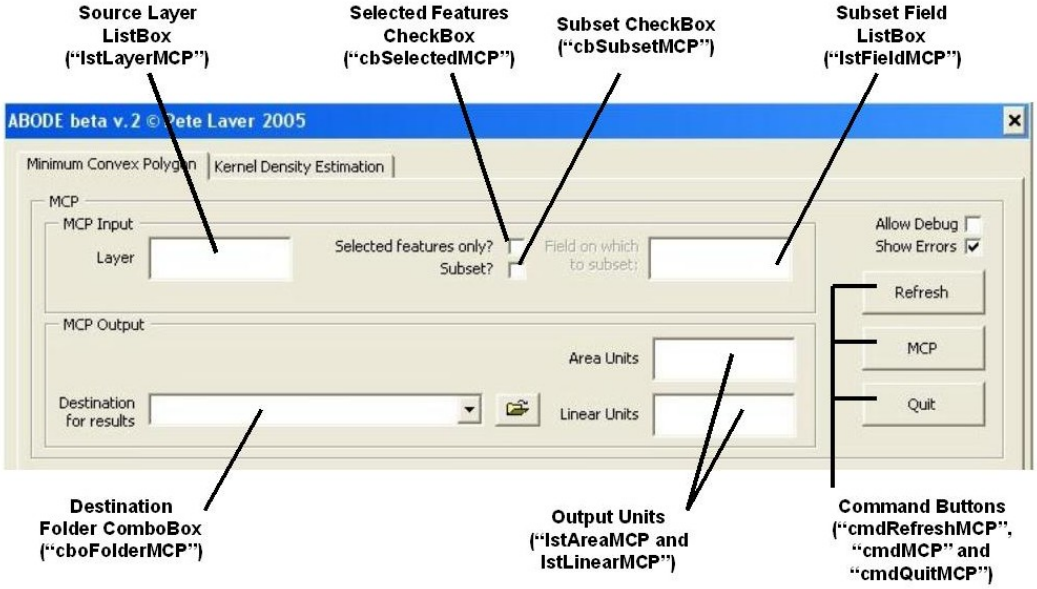


Figure 8.2.2.2.

The code used to run the Minimum Convex Polygon section of this form was adapted with permission from Dr. M. Sawada from the University of Ottawa (Convex Hull). ABODE will automatically register point shapefiles in the table of contents and display the available layers in the Source Layer ListBox (Figure 8.2.2.2.). As soon as the user selects one of the point layers, the form will automatically be 'primed' for that layer. This means that when the user then selects, for example, to subset the dataset (Subset CheckBox, Figure 8.2.2.2.), the available fields from

the selected layer (Source Layer) become activated in the Subset Field ListBox. Once a layer is selected, there is one more required input, the destination for the files that are created. This can be selected from, or entered into, the Destination Folder ComboBox. Choosing to run the analysis on only selected features is an option (Selected Features CheckBox) (features can be selected in ArcMAP using either the Attribute Table or the Select features tool – features are selcted if they are highlighted i.e. in the light blue default color). Please note that if no features in the layer chosen are selected, ABODE will default to using all of the features in the layer. Conversely, if features are selected in ArcMap in the layer, but the Selected Features CheckBox is not checked, then ABODE will again default to using all of the features in the layer. To reiterate, ABODE will run an analysis on a selected feature only if features are in fact selected in a point layer, and if the Selected Features CheckBox is checked. Subsetting on a particular field will provide individual Minimum Convex Polygons for all the unique values contained in that field. This means that all the features that have a common value for that particular field will be grouped together for the analysis. One shapefile for each subset will be created. Choosing selected features works within the subsetting function, and the same rules apply concerning selection requirements (see above). Finally, you can set the output units for the Attribute Tables of the polygons and for any summary tables that may be created by ABODE. If no areal or linear units are selected, ABODE will default to using the projected units from the Map.

Left clicking on the Quit Button (cmdQuitMCP) will remove the form from sight. The Refresh Button (cmdRefreshMCP) will clear the input selections and leave a fresh form for the user. Home range asymptote analyses are commonly done using Minimum Convex Polygons. Please read Section 5 for justification for performing this analysis and for more detail on the theory behind home range asymptotes. The bottom section of the Minimum Convex Polygon Page deals with this analysis (Figure 8.2.2.3.).

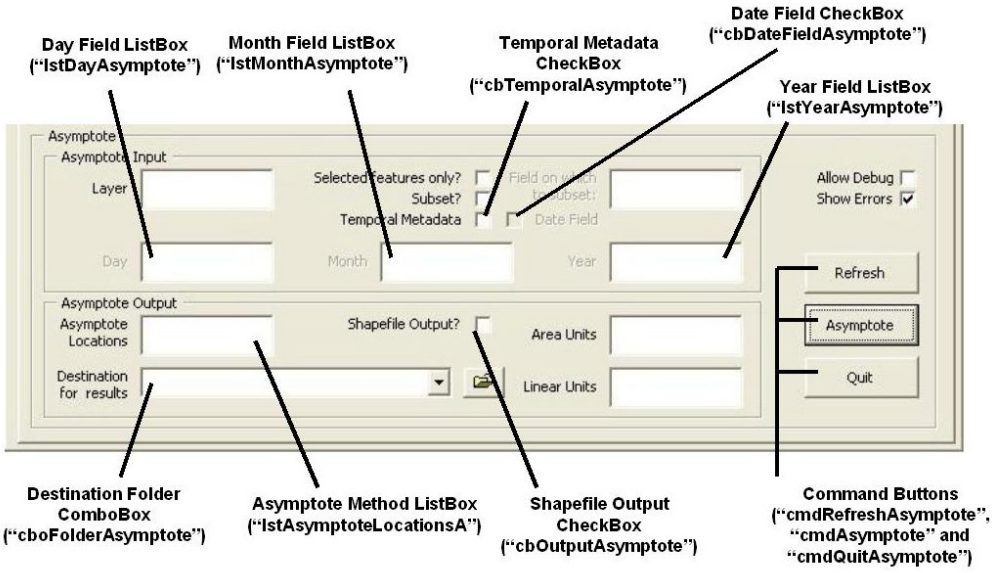


Figure 8.2.2.3.

This analysis runs in the same manner as the regular Minimum Convex Polygon analysis (see above). The difference in this case, however, is the number of points used each time a polygon is created. Initially, just three points are chosen. Then as each new point is added to this subset, a new polygon is created. Once all the points have been used in the analysis, a table is created that contains the number of points and the area for each polygon. This table can then be graphed in a software package such as MS Excel. In ABODE, only the polygons that show an increase in area from the previous polygon subset are displayed. All polygons for which the addition of a point did not increase the polygon size (i.e. these points fell within the perimeter of the previous polygon) are not kept in the destination folder. This reduces the number of superfluous shapefiles stored by ABODE.

The user has the option of keeping the shapefile output from the analysis, which will be displayed in the Table of Contents and in the Map. To do this, check the Shapefile Output CheckBox (Figure 8.2.2.3.). Given a sample size of n points, this analysis could potentially produce n-2 shapefiles (this would occur if every subsequent point occurred outside of the previous polygon boundary). For this reason, the user may wish to leave this checkbox clear. Two methods are available for running the asymptote analysis. The choice of method depends upon the sampling protocol used to collect the data. If the data are discontinuous, then the points should be added to the subset in a random manner; if the data are continuous, then the points should be added sequentially (Harris *et al* 1990). The user will need to select either "Random" or "Consecutive" from the Asymptote Method ListBox for either of the methods (Figure 8.2.2.3.).

An asymptote analysis requires temporal metadata (data such as day. Month and year associated with each location in the dataset), since it runs either on sequential or temporally random points. If no temporal metadata exist, ABODE will assume that the data points are in sequential order in the layer, such that they are ordered indirectly by the field "OID" or the object identifier (an OID is given to each record in the dataset automatically by ArcMAP when data are loaded as a shapefile – these IDs are numbered in order of the records in the table). Metadata (should they be available) are required in the form of separate fields for the day, month and year. Manipulation of a date field in order to obtain this format can be done in MS Excel or MS Access before importing the point data into ArcMap. Once the user checks the Temporal Metadata CheckBox, the Day, Month and Year Field ListBoxes will be activated (Figure 8.2.2.3.). Choose the correct field for each of these temporal categories.

ABODE adds consecutive points by looping through the attribute table, ordered according to the temporal metadata fields selected as inputs. If these are not available, ABODE will loop through the table of the points in the order that they appear in the attribute table (i.e. ordered by the OID). The Random analysis is achieved by adding a field to the table of the points. This field is then filled by random numbers generated by ArcMap. The table is then ordered based on this field, such that the table becomes randomized. The new randomized table is then used in the analysis (where ABODE again loops through the table). Given that no seed value is set for ArcMap's randomization procedure, each time this analysis is run, a new starting subset and new subsequent subsets will be generated. Thus the graph depicting the home range asymptote is no longer deterministic (as is the case for consecutively/sequentially added points), but is now subject to stochasticity (choice of a seed value and randomization of points). For this reason, many iterations of this analysis should be done.

Please note that if an analysis is completed successfully, the user form will be closed automatically. If you were to run another analysis however, the choices and values entered on the form will be retained. If you fill in the inputs for the form, and then left-click "Quit" these inputs will also be retained. However, if your analysis does not run successfully, then the form will be cleared. Likewise, when the user closes the user form manually, the form will be cleared. If a point layer is added to or removed from the ArcMap Table of Contents and no change has been made to the form since it last ran successfully, or since the user clicked "Quit", then the Source Layer ListBox will not register the change. If you left-click on "Refresh" or close the form manually, or if the last analysis was unsuccessful, then the change should register in the Source Layer ListBox. .

8.2.3. Kernel Density Estimation

The Kernel Density Estimation Page consists of three sections (Figure 8.2.3.1.). All three sections are required to run a kernel analysis. These sections group user inputs into the categories of input data, user-defined options, and output data. The top section of the Kernel Density Estimation Page is the Input section (Figure 8.2.3.2.).

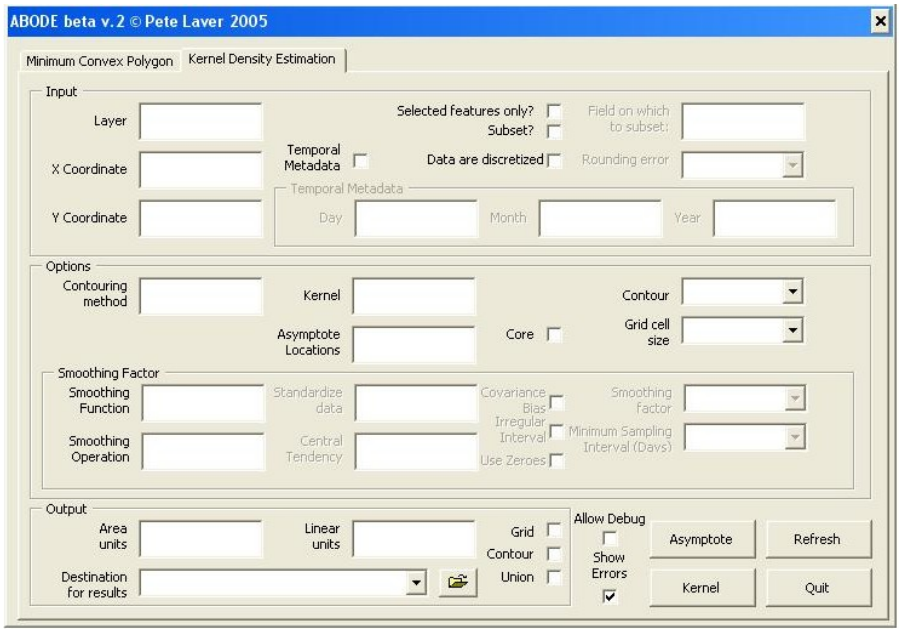


Figure 8.2.3.1.

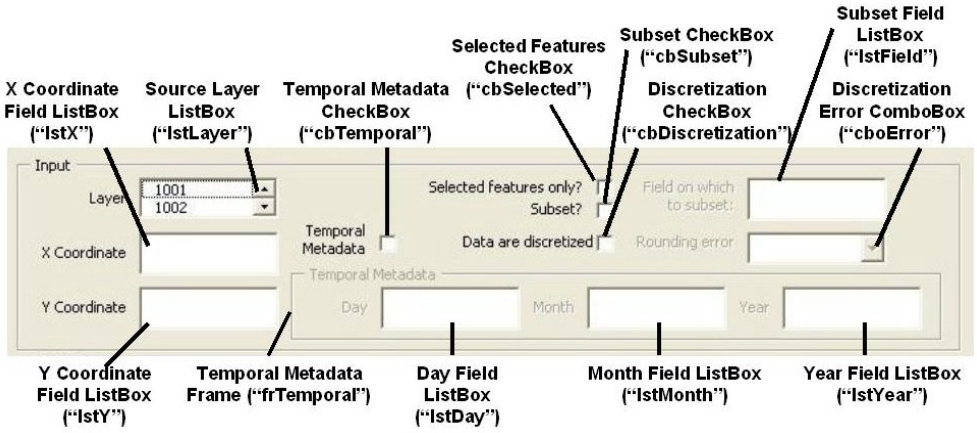


Figure 8.2.3.2.

As with the Minimum Convex Polygon Page, ABODE will search the ArcMap Table of Contents for point layers. It will then display these in the Source Layer ListBox (Figure 8.2.3.2.). For this analysis, the user must indicate which fields contain the coordinate data. Fields that contain data

in a number format will be listed for the selected Source Layer in the X and Y Coordinate Field ListBoxes (Figure 8.2.3.2.). The use of selected features and subsetting follows as for the Minimum Convex Polygon analysis (please see above).

If coordinate data are discretized (i.e. rounded), and if the user has an estimate of the maximum rounding error associated with individual points, then the Discretization CheckBox should be checked (Figure 8.2.3.2.). This option enables the Discretization Error ComboBox, in which the error distance (in map units) should be entered. Please see Section 4.4 for a discussion of this problem. Discretization may result in poor estimation of the required smoothing factor when using Least-Squares Cross-Validation as a method for choosing the smoothing factor (Silverman, 1986; Chiu, 1991; Seaman and Powell, 1996; Kernohan *et al*., 2001). If the Discretization CheckBox is checked, then ABODE will correct for this problem while estimating the smoothing factor using Least Squares Cross Validation. If the distance between two points is calculated as zero map units (i.e. they are at the same place) then the distance will be automatically adjusted to a value selected at random by ArcMap from an even distribution between 1 and the value entered into the Discretization Error ComboBox (Figure 8.2.3.2.) (see Section 4.4). When the actual probability density function is being calculated, then the distances between points will not be adjusted according to this distribution.

Checking the Temporal Metadata CheckBox will enable the Temporal Metadata Frame (Figure 8.2.3.2.). Again the Day, Month and Year Field ListBoxes will register all the fields in the selected Source Layer that have numeric fields. From these options, choose the fields that correspond to the time interval. A kernel analysis may be run without temporal metadata. In order to get the kernel asymptote analysis for random locations (or consecutive where the points are not tabulated in chronological order) and for any analysis involving the “Displacement” smoothing choice (Section 7), these data must be available. Future versions of ABODE will allow the user to select a date field that contains these temporal data. For the present, individual fields are needed for all three. For the use of the “Displacement” smoothing method, the user must indicate what sampling interval was used for data collection (i.e. the time between consecutive locations). If the data were collected using a regular or close to regular schedule, or if the user wishes to manually set the minimum sampling interval, then this value should be entered into or selected from the Minimum Sampling Interval ComboBox. If the interval is not known, or is highly irregular, then the user should check the Irregular Displacement CheckBox. Selecting this option will result in analysis of consecutive points. For each pairing of consecutive points, the time interval in number of days is calculated. The arithmetic mean, or the median, or the mode interval is then estimated. Using either the estimated irregular displacement, or the user defined minimum displacement, ABODE will find all pairs of consecutive points that fall into an interval class that

contains the value. For all of the pairings, the distance between points (in map units) is calculated. Again, an arithmetic mean is generated. This is the mean displacement in distance units per mean sampling interval. This is the displacement distance that will subsequently be used as the smoothing factor. It should be understood that the user has a choice of the mean, median or mode interval between locations, but the distance representing this interval is estimated as only the mean value.

The middle section of the Kernel Density Estimation Page deals with the options that users can select to customize their analysis (Figure 8.2.3.3.).

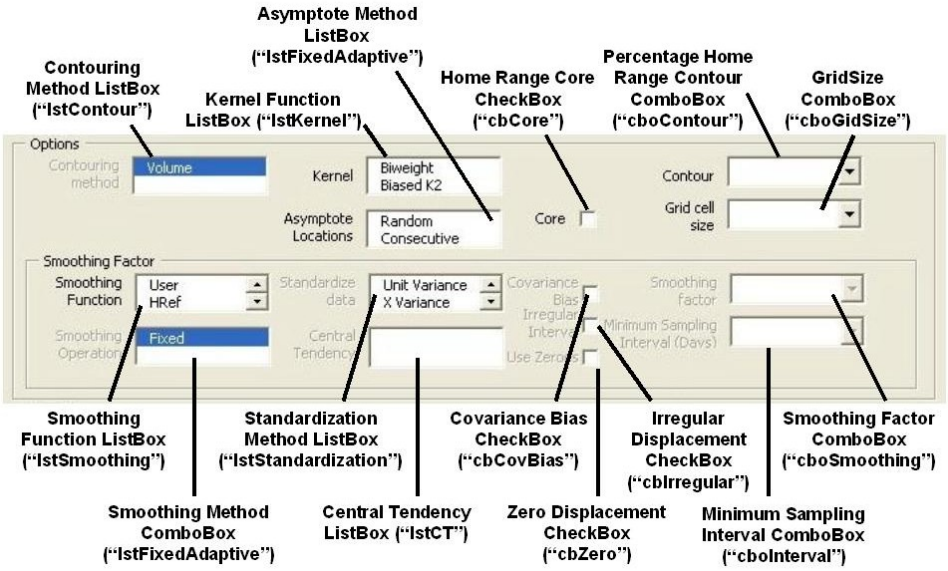


Figure 8.2.3.3.

This section is vital for setting all the required parameters for the kernel analysis. The selections on this part of the page are all required to qualify the result of the analysis when publishing results. The first selection a user should make is whether or not to standardize the data. Data should, in most cases, be standardized (Silverman, 1986; Worton, 1989; Seaman and Powell, 1996; Rodgers and Carr, 1998) (see Section 4.5). Currently, standardization is only available for the Least Squares Cross Validation procedure. When running the analysis using the reference smoothing factor "HRef" then data should be left at their original scale, since the smoothing factor is calculated for the data assuming that the distribution is already bivariate normal (i.e. theoretically, standardization of the data would result in exactly the same distribution as the original data). For user defined smoothing, the smoothing factor is set either arbitrarily, or based on some prior knowledge about the data structure or biological movement behavior of the animal. Thus, it is set disregarding the distribution of the underlying data and should be applied at the

original scale. For Displacement smoothing, the smoothing factor is (as with user defined) a biologically meaningful and data driven value and should be applied as such to the original dataset and not to standardized data. Should Least Squares Cross Validation be chosen as a method for generating a smoothing factor, then the options to use either “Unit Variance” standardization, or “X Variance” standardization become available. Both methods should provide equivalent results (Rodgers and Carr, 1998) though the detail in the contouring may be slightly different. For the other smoothing function options, the user will not be allowed to choose a standardization option, and the option “None” should be chosen. By checking the Covariance Bias CheckBox, the user is able to standardize according to the method for standardization used by Seaman and Powell (1996). This method introduces bias since it is unable to account for covariance in the dataset. Leaving the Box unchecked will allow for standardization to occur in a manner that accounts suitably for covariance (Silverman, 1986:77).

The user then needs to decide on the method for contouring in the Contouring Method ListBox (Figure 8.2.3.3.). Two options are possible for contouring - by density or by volume (Rodgers and Carr, 1998). In ABODE, only the volume contouring option is currently available. This option is used widely for home range analyses (Rodgers and Carr, 1998). This means that when the user wishes to contour at a certain percentage of the home range (e.g. 95%), contouring takes place at that percentage volume of the probability density function and not percentage of the total home range area. In the Kernel Function ListBox, the user can select the form of the kernel used in the density calculation. Silverman (1986) provided the Biweight Kernel (K_2) as an option. This was used by Seaman and Powell (1996) and is the only option in ABODE.

In the Percentage Home Range Contour ComboBox the user can set the required contour level. If the Home Range Core CheckBox is checked, then the options for the percentage contour will be disabled. If this option is chosen, ABODE calculates home ranges for 1%, 5% to 95% (in 5% increments), and 99.9% home ranges. If a core area exists, then ABODE will find the 5% range in which it occurs, and calculate home ranges in 1% percent increments. If a core area exists, it will be output as a shapefile (contoured, to the nearest percentage of home range volume). A table will also be created which contains, for each percentage home range, the percent maximum probability and the percent home range area. These values should be graphed in a graphing package such as MS Excel to check that a concave curve (clumped use) is displayed (see Section 6).

Grid cell size may be set in the GridSize ComboBox. Setting this parameter should be determined by the tradeoff between speed or efficiency of the analysis, and its resolution. The choice should reflect the scale of the data such that the spread in X and Y direction would be covered by an adequate number of cells. A cell size that is very small relative to the scale of the

data will produce a fine resolution for the home range contour. Contours will look smooth and show fine detail. Such analyses will take longer and the resultant grids will require more disk space if kept. Choosing a cell size that is large relative to the scale of the data will result in a quicker analysis, but will be coarse and may not show the detail necessary. At such a scale, the shape and smoothness of the contours is largely governed by the contouring smoothing function in ArcMap. This is the method that ArcMap uses to interpolate between the center point of each cell. ABODE contains options that allow for relative grid sizes - for example, "Resolution 25" would produce a grid cell size based on the size required to fit 25 cells into the smaller of the two ranges for X and Y data. Such options will optimize the tradeoff between resolution and speed.

If you wish to run an asymptote analysis, then the asymptote method (consecutive or random) must be selected in the Asymptote Method Listbox. The procedure for sorting and randomizing the points is equivalent to that used in the MCP asymptote technique (see Section 8.2.2.).

Specific parameters for the smoothing factor (window or bandwidth) may be set in the Smoothing Factor Frame. Firstly, the smoothing Function may be set in the Smoothing Function Listbox. This determines how the smoothing factor is selected. If "LSCV" (Least-Squares Cross-Validation) is chosen, then ABODE will determine the optimal smoothing factor for the given dataset (Section 4.3.2.). This method involves minimization of the mean integrated square error for the density, over various values for the smoothing factor (*h*) (Silverman, 1986). This is achieved through the use of minimizing function GOLDEN (Sprott, 1991). LSCV will in most cases provide the optimal choice of smoothing factor (Seaman and Powell, 1996). Selecting "HRef" (Reference smoothing) will provide the smoothing factor based on the assumption that the given data fit a bivariate normal distribution. With most animal movement distributions, this assumption is violated and this option will thus produce an overestimate of the required smoothing factor and subsequently an overestimate of the home range size. As noted by Silverman (1986), this option may be a good choice for exploratory analyses. Choosing the "User" option will enable the Smoothing Factor ComboBox and allow the user to enter or select a smoothing factor. This choice may be based on prior knowledge of the underlying data structure, or on the knowledge of the behavior of the animal in question. Finally, choosing "Displacement" will provide a smoothing factor that is calculated by ABODE based on the data provided. ABODE takes either the given minimum sampling interval or the calculated (mean, median or mode) sampling interval and estimates the mean distance traveled (crow flight distance) for that time interval. This distance is then used as the smoothing factor. This option will not be robust in all data sets, and is especially sensitive to the sample size, such that enough samples for the selected or estimated time interval are needed to provide a good estimate for the smoothing factor. The theory behind this method suggests that it is both data driven and biologically

motivated. It does however need to be tested extensively, and should be used with caution and a healthy dose of pragmatism.

Finally, in the Smoothing Function Listbox, the method for smoothing may be chosen. Currently, only Fixed smoothing is available as an option following the recommendations of Seaman and Powell (1996). Future versions of ABODE will allow for Adaptive smoothing. The lower section of the Kernel Density Estimation Page allows the user to customize the output provided by ABODE (Figure 8.2.3.4.).

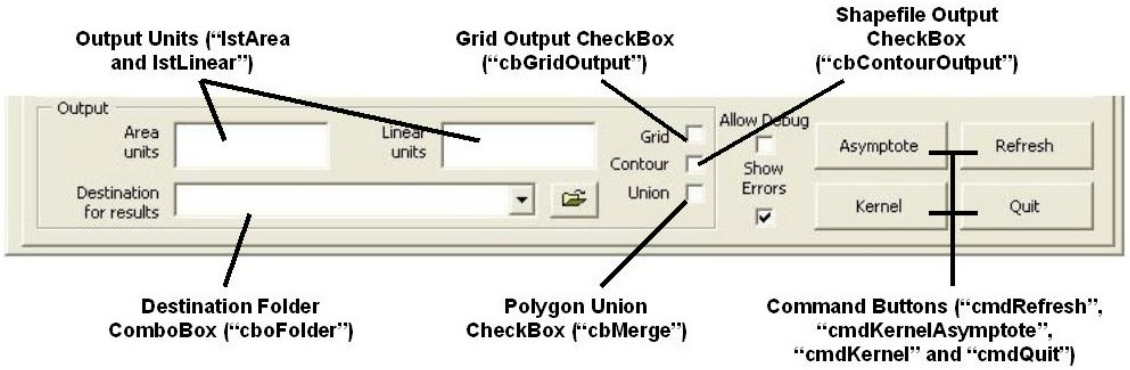


Figure 8.2.3.4.

In the Area and Linear Output ListBoxes, the units of output (in the attribute tables and generated summary tables) can be converted from the map units to the required units. These units will be displayed in the tables, and all values following this will be in the stated units. Checking the Grid and Shapefile Output CheckBoxes will keep the raster and shapefiles generated by analyses. If an asymptote or core home range analysis is performed, then the user is not given the option of keeping the grids. This is to reduce the amount of data stored to the hard drive. The user may however keep the shapefile output for these analyses. This will allow the user, in the case of the asymptote analysis, to see how the home range size increases as points are added. For the core home range analysis, only the core home range will be kept and displayed. If no core exists, no shapefile will be displayed. The Destination Folder ComboBox allows the user to enter or select the folder in which to store the results.

Command buttons work as for the Minimum Convex Polygon Page. "Quit' will remove the form from view, but will retain the selections until the application is closed. "Refresh" will provide a fresh page on which to start a new analysis. Values that are typed into ComboBoxes will not be removed on the refresh function. "Asymptote" will run the kernel asymptote function, and requires that the user select either "Random" or "Consecutive" in the Asymptote Method Listbox in the Options section. To run either a normal kernel analysis or the core home range analysis the

"Kernel" button should be used. Asymptote options will be ignored if the "Kernel" option is chosen. Likewise, the core option will be ignored if the "Asymptote" option is chosen.

9. Conclusion

Be faithful to your data, and to the natural history of your animal. Choose a home range estimator (or inputs) that will best reflect the objective of your study, that will match the underlying data structure, and that will be biologically meaningful (considering the movement behavior of the animal). Always report all user inputs and methodology for the sake of defensibility and replicability.

10. Acknowledgements

Marcella Kelly and the Department of Fisheries and Wildlife Sciences, Virginia Tech, provided financial and logistical support for this project. Marcella helped with the development of theory for alternative deterministic smoothing factor choices. Steve Prisley helped throughout the project, both with intellectual contributions, and with code improvements. Roger Powell provided the impetus for undertaking this project, after many discussions of discrepancies in available software packages, and the inadequacies of current methods in dealing with animal movement. George Terrell provided invaluable help in explaining the intricacies of Kernel Density Estimation. Terrell also provided suggestions for improving the speed of the estimation procedures. Bill Cartstensen provided much ArcObjects/VBA assistance. Dean Stauffer provided intellectual help with ideas about smoothing factors and provided support in elucidating differences between the available software packages throughout the development process. Jay McGhee was a sounding board for the many intellectual and code problems faced during development. Art Rodgers and Angus Carr were helpful in describing methodology for Home Range Extension – after which much of ABODE's functionality was modeled. Finally, most of the code was adapted from snippets and samples provided by ESRI, and especially by ESRI users in the User Forums on www.esri.com, including code that I integrated for generating Minimum Convex Polygons from Dr M. Sawada from the University of Ottawa. Without the availability of code this project would not have been possible.

11. References

Anderson, D.J. 1982. The home range: A new nonparametric estimation technique. Ecology 63:103-112.

Bekoff, M., and Mech, L.D. 1984. Simulation analyses of space use: home range estimates, variability and sample size. Behavior Research Methods, Instruments and Computers 16:32-37.

Bowman, A.W. 1984. An alternative method of cross-validation for the smoothing of density estimates. Biometrika 71(2)353-360.

Burt, W.H. 1943. Territoriality and home range concepts as applied to mammals. Journal of Mammalogy 24:346-352.

Chiu, S.T. 1991. The effect of discretization error on bandwidth selection for kernel density estimation. Biometrika 78:436-441.

Dixon, K.R., and Chapman, J.A. 1980. Harmonic mean measure of animal activity areas. Ecology 61:1040-1044.

Dunn, J., and Gipson, P. 1977. Analysis of radiotelemetry data in studies of home range. Biometrics 33:85-101.

Epanechnikov, V.A. 1969. Nonparametric estimation of a multidimensional probability density. Theory of Probability and its Applications 14:152-158.

Forde, P. 1989. Comparative ecology of muntjac *Muntiacus reevesi* and roe deer *Capreolus capreolus* in a commercial coniferous forest. Ph.D. thesis, University of Bristol.

Gautestad, A.O., and Mysterud, I. 1995. The home range ghost. Oikos 74:195-204.

Gautestad, A.O., and Mysterud, I. 1993. Physical and biological mechanisms in animal movement processes. Journal of Applied Ecology 30:523-535.

Gitzen, R.A., and Millspaugh, J.J. 2003. Comparison of least-squares cross-validation bandwidth options for kernel home-range estimation. Wildlife Society Bulletin 31(3):823-831.

Hansteen, T.L., Andreassen, H.P., and Ims, R.A. 1997. Effects of spatiotemporal scale on autocorrelation and home range estimators. Journal of Wildlife Management 61:280-290.

Harris, S.; Cresswell, W.J.; Forde, P.G.; Trewhella, W.J.; Woollard, T.; and Wray, S. 1990. Home-range analysis using radio-tracking data - a review of problems and techniques particularly as applied to the study of mammals. Mammal Review 20(2/3):97-123.

Hawes, M.L. 1977. Home range, territoriality, and ecological separation in sympatric shrews, *Sorex vagrans* and *Sorex obscurus*. Journal of Mammalogy 58:354-367.

Hayne, D.W. 1949. Calculation of size of home range. Journal of Mammalogy. 30:1-18.

Hooge, P.N. and Eichenlaub, B. 2000. Animal movement extension to Arcview. ver. 2.0. Alaska Science Center - Biological Science Office, U.S. Geological Survey, Anchorage, AK, USA.

Horner, M.A., and Powell, R.A. 1990. Internal structure of home ranges of black bears and analyses of home range overlap. Journal of Mammalogy 71:402-410.

Jenrich, R.I., and Turner, F.B. 1969. Measurement of noncircular home range. Journal of Theoretical Biology 22:227-237.

Kenward, R.E. and Hodder, K.H. 1996. Ranges V: an analysis system for biological location data. Institute of Terrestrial Ecology, Furzebrook Research Station, Wareham, UK.

Kernohan, B.J., Gitzen, R.A., and Millspaugh, J.J. 2001. Analysis of Animal Space Use and Movements. In: Millspaugh, J.J., and Marzluff, J.M.(Eds), Radio Tracking and Animal Populations, Academic Press, San Diego, pp 126-166.

Larkin, R.P., and Halkin, D. 1994. A review of software packages for estimating animal home ranges. Wildlife Society Bulletin 22(2):274 – 287

Laundre, J.W., and Keller, B.L. 1984. Home-range of coyotes: a critical review. Journal of Wildlife Management 48:127-139.

Lawson, E.J.G, and Rodgers, A.R. 1997Diffreneces in home-range size computed in commonly used software programs. Wildlife Society Bulletin 25(3):721-729.

Mitchell, M.S., and Powell, R.A. 2003. Linking fitness landscapes with the behavior and distribution of animals. In: Bissonette, J.A., and Storch, I.(Eds.), Landscape Ecology and Resource Management: Linking Theory with Practice, Island Press, Washington, DC, USA, pp 93-124.

Mohr, C.O. 1947. Table of equivalent populations of North American small mammals. American Midland Naturalist 37:223-249.

Otis, D.L.; and White, G.C. 1999. Autocorrelation of Location Estimates and the Analysis of Radiotracking Data. The Journal of Wildlife Management 63(3):1039-1044.

Peters, R. 1978. Communication, cognitive mapping, and strategy in wolves and hominids. In: Hall, R.L., and Sharp, H.S.(Eds.), Wolf and Man: Evolution in parallel, Academic press, New York, pp 95-108.

Powell, R.A., Zimmerman, J.W., and Seaman, D.E. 1997. Ecology and behavior of North American black bears: home ranges, habitat and social organization. Chapman & Hall, London.

Powell, R.A. 2000. Animal home ranges and territories and home range estimators. In: Boitani, L., and Fuller, T. (Eds.), Research Techniques in Animal Ecology: Controversies and Consequences. Columbia University Press, New York, pp 65-110.

Rodgers, A.R., and Carr, A.P. 1998. HRE: The Home Range Extension for Arcview™: User's Manual. Centre for Northern Forest Ecosystem Research. Ontario Ministry of Natural Resources.

Rudemo, M. 1982. Empirical choice of histograms and kernel density estimators. Scandinavian Journal of Statistics 9:65-78.

Seaman, D.E. 1993. Home range and male reproductive optimization in black bears. Ph.D. thesis, North Carolina State University, Raleigh.

Seaman, D.E.; Millspaugh, J.J.; Kernohan, B.J.; Brundige, G.C.; Raedeke, K.J.; and R.A. Gitzen. 1999. Effects of Sample Size on Kernel Home Range Estimates. Journal of Wildlife Management 63(2): 739-747.

Seaman, D.E.; and Powell, R.A. 1996. An Evaluation of the Accuracy of Kernel Density Estimators for Home Range Analysis. Ecology 77(7):2075-2085.

Silverman, B.W. 1986. Density estimation for statistics and data analysis. Chapman and Hall, London.

Siniff, D.B., and Tester, J.R. 1965. Computer analysis of animal-movement data obtained by telemetry. Bioscience 15:104-108.

Sprott, J.C. 1991. Numerical Recipes Routines And Examples in BASIC. Cambridge University Press, Cambridge, England.

Stickel, L.F. 1954. A comparison of certain methods of measuring ranges of small mammals. Journal of Mammalogy 35:1-15.

Swihart, R.K., and Slade, N.A. 1985. Influence of sampling interval on estimates of home-range size. Journal of Wildlife Management 49:1019-1025.

van Winkle, W. 1975. Comparison of several probabilistic home-range models. Journal of Wildlife Management 39:118-123.

White, G.C., and Garrot, R.A. 1990. Analysis of wildlife radio tracking data. Academic Press, San Diego.

Woodroffe, R., and Ginsberg, J.R. 1998. Edge effects and extinction of populations inside protected areas. Science 280(5372):2126-2128.

Worton, B.J. 1995. Using Monte Carlo simulation to evaluate kernel-based home range estimates. Journal of Wildlife Management 59:794-800.

Worton, B.J. 1989. Kernel Methods for Estimating the Utilization Distribution in Home-Range Studies. Ecology 70(1):164-168.

Worton, B.J. 1987. A review of models of home range for animal movement. Ecological Modeling 38:277-298.